THIS WEEK

EDITORIALS

WORLD VIEW Quicker diagnosis is key to the fight against Ebola outbreak **p.145**

LEVITATION Double-magnet trick manages to float nylon screw **p.146**

INTERRUPTED Amorous frogs prove easier target for hungry bats **p.146**

Genetic rights and wrongs

Australia's decision to uphold a patent on biological material is in danger of hampering the development of diagnostic tests.

It is perhaps unexpected that the United States — the oft-lampooned home of patents on peanut-butter-and-jelly sandwiches and ways to swing a swing — is emerging as one of the most hostile towards patents on naturally occurring genes.

Last week, Australia had the opportunity to join the United States in taking a dim view on such licensing of nature. But a federal court there instead upheld a patent claim on the cancer-associated gene *BRCA1*. In doing so, the country remains with Canada, Japan and several countries in the European Union, all of which, unlike the United States, recognize such patents.

The patent on *BRCA1* has become a touchstone in the debate over 'gene patents', a broad term that can cover a wide swath of patent claims on DNA sequences. Certain mutations in *BRCA1* increase the risk of, in particular, breast and ovarian cancers. And Myriad Genetics, a genetic-testing company in Salt Lake City, Utah, has aggressively defended its patents, which cover the abnormal *BRCA1* sequence and tests to identify it.

In the United States, debate on gene patenting has been tied to clear public-health concerns. Myriad's monopoly bred worry that women would have only a single option for *BRCA1* testing, with no possibility of receiving a second, confirmatory test elsewhere. So when advocates challenged US patents on *BRCA1* and the closely related gene, *BRCA2*, the case provoked a passionate response from the public. The patents were defeated in a landmark decision last year that changed decades of legal practice in the field (see *Nature* **498,** 281–282; 2013).

Australia is in a different situation. The *BRCA* patents have not been enforced there, either by Myriad or by the company that has licensed them in Australia: Genetic Technologies of Melbourne. Despite the fervent involvement in the case of patient advocates, including cancer survivors, the spectre of gene patents in Australia remains more theoretical.

Still, the case, and the attention it has received, shines new light on how such patents will affect the future of medical diagnostics, particularly as genetic tests expand to cover large numbers of genes, and even the full genome.

In the run-up to the US decision on Myriad, an academic subfield was born from a handful of patent lawyers and scholars who wanted to know just how many gene patents might be affected — no easy task in a system that awards around 300,000 patents a year. Answers varied, but the general conclusion was: not as many as you might think. One study found that most patents that mention DNA sequences do not claim the sequences as an invention (O. A. Jefferson *et al. Nature Biotechnol.* **31,** 1086–1093; 2013). Nevertheless, even a few hundred patents on genes can be enough to scare off potential investors and entrepreneurs looking to pioneer methods of genetic testing, because they might infringe on protected genetic property.

In the United States, gene patents were defeated because they ran afoul of a prohibition on patents claiming a "product of nature". An influential brief to the US Supreme Court written by biologist Eric Lander, head of the Broad Institute in Cambridge, Massachusetts, wounded the held idea that isolating DNA changed it from a product of nature to a human product, thereby making it patentable. He pointed out that 'isolated' bits of DNA can be found floating free in the blood, and that the isolated *BRCA1* and *BRCA2* genes had in fact been found doing just that.

In Australia, no such limitation on patenting natural products exists. Instead, the debate there has centred on whether the patent claims a "method of manufacture". Last week, five Australian justices unanimously ruled that it does, because to isolate DNA from its natural setting requires effort. This, they say, describes "an artificially created state of affairs for economic benefit", and is therefore fair game for a patent.

> ***"Debate on gene patenting has been tied to clear public-health concerns."***

The plaintiffs in the case are considering an appeal to the High Court of Australia, and some patient advocates are crying for changes to the law to do away with gene patents.

For now, the Australian decision is certain to please patent lawyers and some biotechnology executives. This seems to have been the justices' intent: the bulk of their rulings have focused on preserving incentives for innovation and business. There has been little, if any, attention from the court to what this means for science, or for patient access to information about their genes.

Business concerns are important: the biotechnology industry depends on patents for its livelihood, and many patients' lives depend on the industry. But the business model pursued by Myriad is a fading one, and it is time to look to the future. That future has little place for patents that could hold up the development of bigger and better medical tests. ■

Ebola: time to act

Governments and research organizations must mobilize to end the West African outbreak.

After disproportionate media attention on Ebola's negligible risk to people in Western and Asian countries, the focus seems at last to be shifting towards how to stop the outbreak in West Africa. The grim reality is that medical organizations are struggling: the flood of new cases far outpaces available beds and treatment centres. Many of those who are ill are not receiving the basic health care that could keep them alive.

The tragedy is that we know how to stop Ebola. Well-informed communities can reduce the main routes of spread by avoiding

unprotected home-based care of infected people and by modifying traditional burial practices. Infection-control measures protect health-care workers. Together with rapid identification and isolation of ill people, and tracing and monitoring of their contacts for 21 days (the maximum incubation period of the disease), such measures have stopped Ebola outbreaks in the past.

But the dysfunctional health-care infrastructure of the three countries at the centre of the outbreak — Guinea, Sierra Leone and Liberia, which are poor and struggling to emerge from years of war — is simply not up to the task. The nations need help, and urgently.

The international community must mobilize now. Aid is increasing, but most of those involved, from governments and the World Health Organization (WHO) to researchers, all initially underestimated the threat. This is perhaps because most past outbreaks have been small and relatively straightforward to control.

The WHO has a part to play, but contrary to a widespread assumption, it does not have the in-house capacity to send large teams into the field. The agency's funding for outbreak responses has been slashed, and it has shifted focus to helping countries to reinforce their health systems so that they can respond better themselves. How the international community should best react to outbreaks, and what role the bureaucratic WHO should have, is a debate for after this outbreak is over. The pressing need now is to bring all available resources and talent to bear.

It is a sign of how desperate the situation has become that on 2 September, Joanne Liu, international president of medical group Médecins Sans Frontières (or Doctors Without Borders), called on countries to immediately deploy their military and civilian biodefence teams — units that have been developed to respond to bioterror attacks. The crucial priorities, she said, are to scale up isolation centres, deploy mobile diagnostic labs (see page 145), build a network of field hospitals and establish dedicated air links to shift staff and equipment to where they are needed. In short, a military-style response, with its associated strong chain of command, logistical capacities and speed. The concept makes a lot of sense and is an approach that governments should consider adopting — or explain why, if they choose not to do so. US President Barack Obama indicated last weekend that he would deploy the US military to assist in the outbreak.

It cannot be repeated enough that public-health measures and good old-fashioned epidemiological tracking of the infected and their contacts will bring this outbreak to an end. The priority must be to scale these up, alongside establishing more Ebola treatment centres on the ground. For instance, Ebola's high death rate could be slashed by better patient care, in particular by giving intravenous rehydration.

“The pressing need now is to bring all available resources and talent to bear.”

A highly effective Ebola vaccine would be a game-changer. A WHO-convened meeting on 4–5 September agreed on an unprecedented set of measures, including relaxing regulatory requirements so that experimental drugs and vaccines can be quickly tested under the difficult field conditions of this outbreak, and perhaps even widely deployed. The measures will, for example, permit expedited vaccine trials and informal clinical studies of drugs that could produce useful initial data within months.

Regulators and researchers should be applauded for their speed and pragmatism in exploring innovative methods for conducting trials during this outbreak. Crucially, all those who organize trials must be willing to standardize and share the data they collect to maximize their scientific and medical value, and to allow rapid decisions to be made on which products to prioritize.

West Africa's outbreak illustrates the serious weaknesses in the international community's ability to respond to outbreaks of emerging diseases, despite years of debate. It should also hammer home a truism for future planning — the costs of setting up infrastructure to ensure an early response are small compared with the huge social and economic costs of a large deadly disease outbreak. ■

Orbital assembly

The space launch of a 3D printer does not herald a brave new era — but it is a good start.

Perhaps the most famous DIY job ever done in outer space was performed in April 1970 after an explosion disabled *Apollo 13* on its way to the Moon. The three astronauts on board the craft scrambled together a makeshift adapter from cardboard, plastic bags and duct tape to scrub poisonous carbon dioxide from the air.

What if they had had access to a device that could design and manufacture replacement parts to order?

Last year, an engineer demonstrated just such a device: a three-dimensional (3D) printer. Working for Made in Space in Moffett Field, California, the engineer spent an hour on the computer designing an adapter for the *Apollo 13* scrubber, and the rest of the day printing it and demonstrating how it would work. Problem solved?

Perhaps it would be if every spacecraft had a 3D printer. Working with NASA, Made in Space is about to launch the first such printer into space (see page 156). If dreamers have their way, it will be the start of a new generation of manufacturing in orbit.

Imagine a rocket carrying little but a machine that can print the infrastructure for a colony on Mars. Or a spacecraft that can unfurl robotic tethers, printing and braiding giant ribbons into a starshade so that a telescope can stare, unblinded, at extraterrestrial worlds.

If this sounds impractical, it's because it is. For decades, enthusiasts have dreamed up ambitious ways to manufacture structures in space. A 1970s concept known as the beam builder would have welded aluminium tubes together to create huge trusses spilling out of the back of the space shuttle. But in the 1990s, when countries began building hardware components for the International Space Station, they opted to do so in the familiar environment of planet Earth. Each large element was sent into space individually; only once aloft were the parts joined together to form the sprawling complex.

It would be wise to remember such lessons as the enthusiasm for 3D printing runs high. In July, a US National Research Council report concluded that there is a vast gap between what people think the technology can do and what it really can. It is all very well to pack a 3D printer for a journey to deep space — but what should a space traveller do when the printer itself breaks down? Carry a backup?

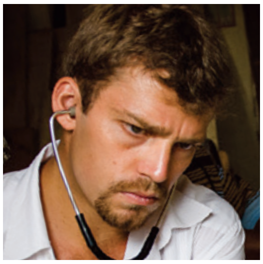
There is a place for 3D printing in space applications. Among other things, designers on the ground can dream up bizarre and fanciful parts, then print them regardless of many conventional design constraints. In principle, this means slimmer spacecraft that are cheaper to launch. That can be a big deal for an industry that must weigh every nut and bolt.

NASA is even talking about printing CubeSats, small box-shaped satellites that can be launched in flocks from a single launch vehicle or off the space station itself. A PrintSat, a CubeSat printed from a polyamide-based material, is slated for launch later this year as a test for how well such devices might survive in the harsh environment of space.

NASA is not alone. The European Space Agency is developing ways to use plastic and metal printed parts on the space station; the Italian Space Agency is hoping to send its own printer to the station in 2017.

Such experiments may not lead directly to a Martian base, but that is no reason not to encourage the fledgling technology. The maker ethos has permeated everywhere, it seems — even beyond the gravitational pull of Earth. ■

NATURE.COM
To comment online, click on Editorials at:
go.nature.com/xhunqv



SARAH BONES

Make diagnostic centres a priority for Ebola crisis

Bottlenecks in testing samples for Ebola leave patients stranded for days in isolation wards and raise fears of seeking treatment, says **J. Daniel Kelly**.

I will never forget the first time I walked into an Ebola isolation ward at Connaught Hospital in Freetown, Sierra Leone. It was 20 August. Inside, eight people thought to have the disease were organized into three patient-care rooms. Patients in the first room appeared to be healthy, and we greeted each other.

In the second room, patients barely had the strength to sit. Still, they were able to articulate how they felt. In the last room there were two patients. One was a woman who seemed confused and agitated, and was later confirmed to have the disease. On the other side of the room, a young man was curled into the corner of his bed. He seemed healthy but was terrified.

He had been deathly ill when he was admitted three days earlier. He recovered, but had watched Ebola kill two others in that room.

I could only imagine how I would feel in that situation, watching others get sick and die, wondering if I would be next. Then I considered the deplorable conditions — no visitors were allowed, and a bucket served as a bathroom — and how I, wearing my protective 'spacesuit', must have looked to the curled man. The idea of becoming sick with Ebola in Sierra Leone frightened me.

It frightened him too, and much of his fear could have been avoided. It took four days for his blood to be tested and shown to be free of Ebola. At that point, Sierra Leone had two facilities able to diagnose the virus. The nearest — Kenema Government Hospital — was five hours away and was overloaded with blood samples from around the country.

The evening the curled man arrived at Connaught, there was no nursing staff to oversee patient care. The Sierra Leonean doctor who had supervised the ward had died, and no Sierra Leonean doctor had taken his place. The man was locked in this terrifying environment until someone could draw his blood for testing. Blood samples and sick patients were sent to Kenema by ambulance only at the end of each day. Even after the man's blood sample arrived in Kenema, it was not tested until the next day.

I began working in Sierra Leone eight years ago, when I co-founded Wellbody Alliance, a non-profit health-care organization in Kono, so I am familiar with the logistical challenges facing the country's collapsing health-care system. But the desperate shortage of Ebola diagnostic centres in Sierra Leone is fuelling the Ebola outbreak. People who think that they might have the disease do not want to spend several days trapped in an isolation unit, away from their families and surrounded by workers in spacesuits.

This fear means that patients go to isolation wards only when their symptoms are severe, if they go at all. If Sierra Leone's Ministry of Health and Sanitation could scale up diagnostic facilities, it would reduce fear and help to curb transmission from very sick people who are reluctant to seek treatment.

Take Freetown, for example. A four-person team from South Africa arrived there on the same flight that I did. They came with a machine for analysing viral RNA and created a diagnostic site in the outskirts of Freetown at the National Laboratory of Sierra Leone. Within a week, the team was sending Ebola test results to the isolation ward twice a day. Some patients did not even have to stay overnight. That kind of experience feels less daunting and more acceptable.

Even though Freetown now has a faster turnaround time on test results, Port Loko, the latest Ebola hot spot, is still sending blood samples to Kenema. In Kono, where I have also visited, three patients had to wait for their blood sample results to come back from Kenema for confirmation of diagnosis. The delay meant that all three died before they could be transferred to a treatment centre.

> PATIENTS GO TO **ISOLATION** WARDS ONLY WHEN THEIR SYMPTOMS ARE **SEVERE,** IF THEY **GO AT ALL.**

Two weeks ago, Tom Frieden, director of the US Centers for Disease Control and Prevention (CDC), warned that it was only a matter of time before the Ebola outbreak in Sierra Leone would escalate to match the situation in Liberia. The World Health Organization (WHO) and other modelling experts have predicted 20,000–100,000 Ebola infections before the epidemic is over. We need to minimize delays in care and if we cannot speed up the health system's lethargy, then we need to bring diagnostics closer to the people. That means we need more diagnostic sites. So far, all such sites have been developed as adjunctive services to treatment centres. We need to expand these services to every district, even those that have only an isolation centre.

Because most of the clinical-care focus has been on isolation and treatment centres, the strategy for diagnostic sites has been overlooked.

One of the challenges is the need to standardize equipment, techniques and results. The Ministry of Health and Sanitation wants standard diagnostics, and international agencies such as the CDC and the WHO agree. Standardization takes time, but it is necessary. Sierra Leone uses at least four different types of donated protective suit in its isolation wards, which can change the decontamination process and confuse health workers.

As the number of suspected Ebola infections in Sierra Leone rises, its health system will be under increasing strain to deliver test results in a timely fashion. Three diagnostic sites are not enough. ■

NATURE.COM
Discuss this article online at:
go.nature.com/8qnftb

J. Daniel Kelly *is an infectious-disease fellow at the University of California, San Francisco.*
e-mail: dan.kelly@ucsf.edu

RESEARCH HIGHLIGHTS

Selections from the scientific literature

GENOMICS

How coffee got its buzz

The coffee plant makes caffeine using different genes from those found in tea and cacao, suggesting that the ability to produce the stimulant evolved at least twice in plants.

Victor Albert at the University of Buffalo in New York and his colleagues sequenced the genome of robusta coffee, *Coffea canephora*, which makes up about one-third of all coffee produced. They found that of the genes that are unique to this plant, most are involved in caffeine production.

The stimulant probably evolved in the ancestor of coffee plants and separately in a common ancestor of tea and cacao, perhaps to defend the plants against predators and to attract pollinators.

Science **345, 1181–1184 (2014)**

TINA GUNHOLD

ANIMAL BEHAVIOUR

Videos teach tricks to wild monkeys

Wild monkeys can learn new behaviours by watching instructional videos — a feat that had previously been accomplished only in the laboratory.

Tina Gunhold at the University of Vienna and her collaborators recorded video of two captive marmosets (*Callithrix jacchus*) as they opened a box to retrieve a reward, either by popping open a lid or by pulling out a drawer. The team then placed the same type of box on a tree branch in a Brazilian forest, next to a laptop screen showing the videos (pictured).

Of the 108 wild animals studied, only 12 succeeded in the task, but 11 of those had watched the video. Most of the successful animals used the same technique they had seen in the video.

Biol. Lett. **10, 20140439 (2014)**

EVOLUTION

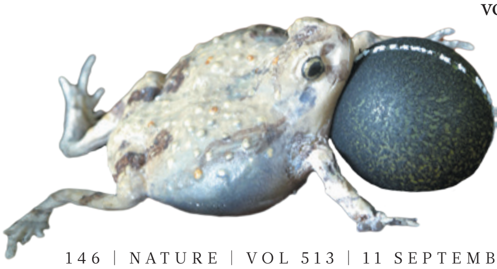
Wooing frogs are bat bait

Bats use echolocation not only to navigate, but also to spot and capture male frogs that are in the act of courting.

Many male frogs inflate their vocal sacs while sending out calls to attract potential mates. Wouter Halfwerk at the Smithsonian Tropical Research Institute in Balboa, Panama, and his team exposed wild-caught fringe-lipped bats (*Trachops cirrhosus*) to robotic models of the male túngara frog (*Physalaemus pustulosus*) that either puffed out a vocal sac (**pictured**) in time with a call or just emitted the call.

They found that all the bats preferentially attacked the model that inflated its sac in sync with the call. The bats used echolocation to detect the 'frogs' from 3–5 metres away, whereas female frogs use vision to assess the male's vocal sac. The results suggest that sexual and natural selection can act on the same trait through different senses.

J. Exp. Biol. **217, 3038–3044 (2014)**



RYAN TAYLOR

PHYSICS

Magnets used in suspension act

Researchers have developed a way to handle small objects in three dimensions using magnetic levitation, even when the objects themselves are not magnetic.

George Whitesides and his team at Harvard University in Cambridge, Massachusetts, suspended a non-magnetic nylon screw in a liquid that becomes magnetic when exposed to magnets. The authors placed one magnet above the container and one below, which made the fluid shift towards the magnets, leaving the screw suspended in the middle.

When the apparatus is rotated, the object rotates with it. Moving an extra magnet around the outside of the device further shifts the liquid and the screw's orientation.

The technique could be useful on assembly lines, allowing the manipulation of materials that are too fragile or soft to be handled by other equipment.

Proc. Natl Acad. Sci. USA **http://doi.org/vgq (2014)**

MARINE ECOLOGY

Blue whales bounce back

A population of blue whales has reached pre-whaling levels and is no longer endangered.

Cole Monnahan at the University of Washington in Seattle and his colleagues modelled a population of blue whales (*Balaenoptera musculus*) in the eastern North Pacific along with the number of ships and their collisions with the mammals between 1905 and 2050. They found that whale numbers in this region were at their lowest in 1931 and have since increased to about 2,200 — nearly the maximum population size that the ecosystem can sustain.

The team also estimates that ship strikes are unlikely to threaten the population in the near future, but says that collision numbers are currently above legal US levels.

Mar. Mammal Sci. **http://doi.org/vh8 (2014)**

MICROBIOLOGY AND IMMUNOLOGY

Early diet shapes gut flora

Breast- and bottle-fed monkeys develop distinct immune systems and communities of gut microbes.

Populations of gut flora vary among adult primates, but little is known about what drives these differences. Dennis Hartigan-O'Connor of the University of California, Davis, and his colleagues found that breast-fed rhesus macaques (*Macaca mulatta*) reared by their mothers had different gut flora from bottle-fed macaques raised in a nursery.

Breast-fed infants also developed a larger population of immune-system cells called T_H17 cells, which are important mediators of anti-pathogen responses. These differences persisted for six months after the infants began receiving identical diets.

Some metabolites, including arachidonic acid, in the macaques' stool correlated with these differences, suggesting that these compounds could be mediating the effects on gut flora and the immune system.

Sci. Transl. Med. 6, **252ra120 (2014)**

NEUROSCIENCE

Music training aids speech processing

The more music training children receive, the better their brains become at distinguishing between similar speech sounds.

Nina Kraus at Northwestern University in Evanston, Illinois, and her colleagues studied children aged six to nine years from low-income neighbourhoods in Los Angeles, California, who took part in an after-school programme of musical instruction. The authors found that children who were in the programme for two years had faster and more-sensitive brainwave responses to syllables such as 'ba' and 'ga' than those who had been enrolled in the class for only a year.

This kind of speech processing is important for reading and language skills, the authors say, adding that music training could improve brain function in children.

J. Neuro. 34, **11913–11918 (2014)**

BIOFUELS

Bacteria generate propane gas

Genetically engineered bacteria could one day be harnessed to make renewable propane fuel.

Patrik Jones at Imperial College London, Kalim Akhtar at University College London and their colleagues introduced genes for various enzymes from different species of bacteria into *Escherichia coli*, so that the microbe could convert glucose into propane gas. With genetic tinkering and by increasing the levels of oxygen to which the engineered bacterium was exposed, the team boosted propane production by two orders of magnitude.

Propane is an ideal biofuel because as a gas, it can be separated from the cultivation medium and easily liquefied for efficient storage, the authors say.

Nature Commun. 5, **4731 (2014)**

SOCIAL SELECTION *Popular articles on social media*

The language of deception

A *PLoS ONE* paper on language patterns in fraudulent papers has sparked social-media speculation about new ways to spot dishonest work. Researchers at Cornell University in Ithaca, New York, took advantage of a singular resource to study the linguistics of fraud: the collected works of Diederik Stapel, a Dutch social psychologist who confessed to faking data in many of his papers. The Cornell team analysed papers that had been deemed fraudulent by three investigative committees, and compared them with his genuine publications. They found that the falsified papers had a linguistic signature. Among other things, they tended to have fewer qualifying words (such as 'possibly') and more amplifying words such as 'extremely'. "Lucky he had enough false papers for analysis!" tweeted Grace Lindsay, a neuroscience graduate student at Columbia University in New York City.

PLoS ONE 9, **e105937 (2014)**

Based on data from altmetric.com. Altmetric is supported by Macmillan Science and Education, which owns Nature Publishing Group.

➲ NATURE.COM
For more on popular papers:
go.nature.com/uzywif

ZOOLOGY

Archer fish show how to sharpshoot

STEPHEN DALTON/MINDEN PICTURES/CORBIS

Archer fish can control the water jets they shoot from their mouths to nab prey from a variety of distances.

Peggy Gerullis and Stefan Schuster at the University of Bayreuth in Germany trained the fish (*Toxotes jaculatrix*; **pictured**) to fire at specific targets from defined locations. They then used a high-speed camera to film the animals as they shot at targets of different heights. They found that the fish adjusted the jets of water so that they were most focused and forceful just before reaching the target. For a target 60 centimetres above the water, the fish produced a jet that remained stable over a longer period of time — by opening its mouth more gradually — than when aiming for a target 20 cm above the water.

This ability is analogous to throwing in humans and could similarly have contributed to the evolution of cognitive skills in the fish, the authors say.

Curr. Biol. **http://dx.doi.org/10.1016/j.cub.2014.07.059 (2014)**

➲ NATURE.COM
For the latest research published by *Nature* visit:
www.nature.com/latestresearch

SEVEN DAYS *The news in brief*

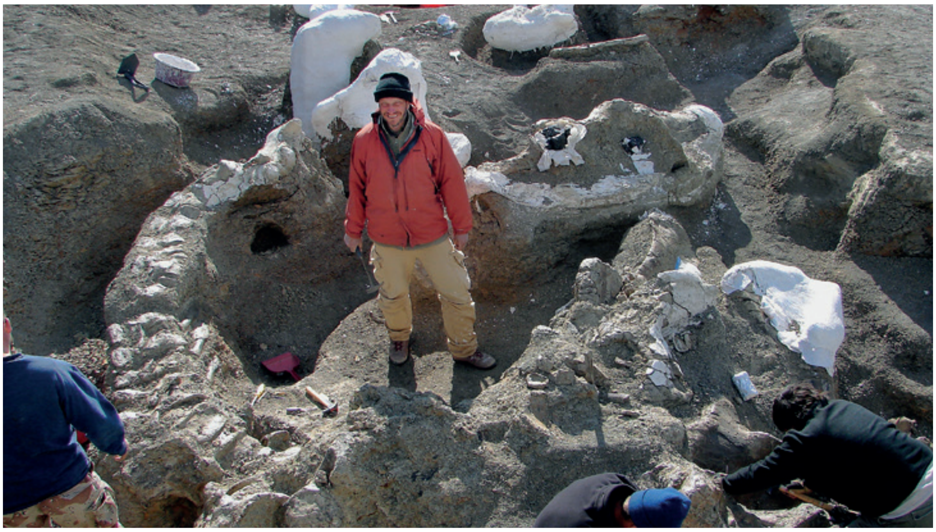
RESEARCH

Quantum chips

Google aims to build its own quantum-computer chips, through a hardware initiative announced on 2 September. The firm, based in Mountain View, California, has partnered on the project with a quantum-computing group at the University of California, Santa Barbara, led by John Martinis. Alongside its own effort, Google says that it will continue to work with computer firm D-Wave in Burnaby, Canada, which sold what was claimed to be the second ever commercial quantum computer to a Google-led collaboration in May 2013 (see *Nature* http://doi.org/mt2; 2014).

Geckos die in space

All five geckos involved in a Russian space-agency sex experiment have died. The animals were launched into orbit in July so that researchers could study the effects of microgravity on animal mating habits. Researchers had feared that the geckos were lost when the spacecraft the animals were in briefly lost communication with ground control (see go.nature.com/iqrbks). The Foton-M4 craft landed safely last week, and although a group of fruit flies had survived, the geckos had perished, the agency says.



KENNETH LACOVARA

South America yields titanosaur bones

Palaeontologists have identified a new species of dinosaur from bones uncovered in southwestern Patagonia, Argentina. Called *Dreadnoughtus schrani*, the animal belongs to a subgroup of sauropods — large, long-necked, herbivores. Kenneth Lacovara (pictured) at Drexel University in Philadelphia, Pennsylvania, led the study, published on 4 September, that excavated two relatively complete specimens of different sizes (K. J. Lacovara *et al. Sci. Rep.* **4,** 6196; 2014). The bones were buried in rocks formed from flood-plain sediments laid down between 66 million years and 84 million years ago. Analysis suggests that the dinosaur stretched about 26 metres from snout to tail. The larger of the specimens, which may not have been fully grown, is estimated to have weighed nearly 60 tonnes. See go.nature.com/7globi for more.

EVENTS

Ebola strategy

The Ebola outbreak in West Africa has claimed more than 2,000 lives, the World Health Organization (WHO) estimated on 5 September, and has potentially infected almost 4,000 people. A wave of cases in Nigeria was traced to an infected traveller in July, and is feared to have spread widely through a doctor who became infected. WHO advisers decided last week that the best experimental treatment to use is a blood transfusion from survivors of the disease. See go.nature.com/m55ual for more.

Ricin left behind

A laboratory sweep at the US National Institutes of Health uncovered forgotten stores of the toxin ricin and four pathogens, according to an agency memo released on 5 September. The agency performed the search after finding improperly stored vials of the deadly smallpox virus in a refrigerator at its campus in Bethesda, Maryland, in July. See go.nature.com/eyyylw for more.

Greenhouse gases

Atmospheric greenhouse-gas levels reached a record high in 2013, the World Meteorological Organization in Geneva, Switzerland, reported on 9 September. The global carbon dioxide concentration hit 396 parts per million (p.p.m.) — a 2.9 p.p.m. increase from 2012, and the largest annual rise since 1984. At that rate, the global CO_2 concentration is set to exceed the symbolic 400 p.p.m. threshold in 2015 or 2016 (see *Nature* **497,** 13–14; 2013).

BUSINESS

Melanoma drug

On 4 September, US regulators issued their first approval of a drug that helps the immune system to fight cancer by blocking a protein called PD-1 (see *Nature* **508,** 24–26; 2014). The drug, Keytruda (pembrolizumab), made by Merck of Whitehouse Station, New Jersey, was granted accelerated approval by the US Food and Drug Administration for patients with advanced melanoma that does not respond to other treatments. Merck plans to charge about US$12,500 for a month's supply of Keytruda.

Gene-patent debate

On 5 September, an Australian federal court dismissed a lawsuit challenging a patent on the gene *BRCA1*. The patent, held by Myriad Genetics of Salt Lake City, Utah, protects

genetic tests for *BRCA1* mutations associated with certain cancers. The decision, which contrasts with that on a US case last year, will fuel the debate about whether naturally occurring genes can be patented. See go.nature.com/ewc9zn and page 143 for more.

FUNDING

Keeling curve cash

The iconic 'Keeling curve', a 56-year-old record of rising atmospheric carbon dioxide levels at Mauna Loa in Hawaii, has won a five-year, US$500,000 grant from US philanthropists Wendy and Eric Schmidt, announced on 3 September. The CO_2-monitoring project, which is maintained by geochemist Ralph Keeling at the Scripps Institution of Oceanography in La Jolla, California, has been threatened by recent funding cutbacks at US government science agencies. See go.nature.com/19a3hv for more.

PEOPLE

Telescope leader

GMTO

Physicist Edward Moses (**pictured**) has been chosen to lead the Giant Magellan Telescope project, the international collaboration announced on 3 September. Scheduled for completion early next decade, the 25-metre telescope in the Chilean Andes will be used to study the formation of stars and galaxies in the early Universe, among other aspects of deep space. Moses formerly served as principal associate director of the Lawrence Livermore National Laboratory in California.

AWARDS

Lasker awards

Kazutoshi Mori of Kyoto University in Japan and Peter Walter of the University of California, San Francisco, have won this year's Albert Lasker Basic Medical Research Award. The prize recognizes their work on how cells correct proteins that are improperly folded. Geneticist Mary-Claire King at the University of Washington in Seattle won a special achievement award for her work on the *BRCA1* gene, which is linked to breast cancer (see go.nature.com/f9miaa for more). Winners of the awards, announced this year on 8 September, often go on to get a Nobel prize.

POLICY

Animal-studies poll

Nearly one in four people believes that the British government should ban all animal research, according to a survey published by the UK Department for Business, Innovation & Skills on 4 September. However, more than two-thirds of the 969 adults surveyed said that it is acceptable "so long as it is for medical research purposes and there is no alternative". Only one-fifth of participants felt that organizations that conduct animal research are well-regulated; the most common perception — reported by 44% of people — was that such organizations are "secretive".

Gender equality

Europe has seen a large increase in the number of nations with quotas or targets for gender equality in public-research leadership positions — from 8 countries in 2008 to 18 in 2013 — according to a European Commission report released on 3 September. The survey of 31 countries found, however, only a slight increase — from 12 to 15 in the same period — in the number of countries whose research institutions had implemented broader gender-equality plans.

Uranium for India

India signed a nuclear-cooperation deal on 5 September to buy uranium for power generation from Australia — the first nation to buy Australian uranium without having signed the international nuclear non-proliferation treaty. Australia also agreed to increase sales of conventional fuels to India, such as coal and natural gas. About one-quarter of India's population of 1.2 billion lacks access to electricity, according to the World Bank.

Aircraft emissions

The US Environmental Protection Agency (EPA) will formally consider whether carbon dioxide produced by aviation poses a threat to human health — the first step towards developing regulations that could restrict aircraft CO_2 emissions. In a 4 September filing to the International Civil Aviation Organization, the EPA said that it will release its findings by 2015, and finalize them in 2016. The agency already limits CO_2 from automobiles and some power plants.

Clean-coal permits

The US Environmental Protection Agency announced on 2 September that it had approved permits for the FutureGen programme to inject carbon dioxide deep underground — a key step in the carbon capture and sequestration demonstration project, which has started and stalled several times since 2003 (see *Nature* **459,** 901; 2009). The project involves capturing CO_2 from a retrofitted, coal-fired power plant in Meredosia, Illinois, and injecting it underground through a series of wells. The permits would allow well-drilling to begin next month.

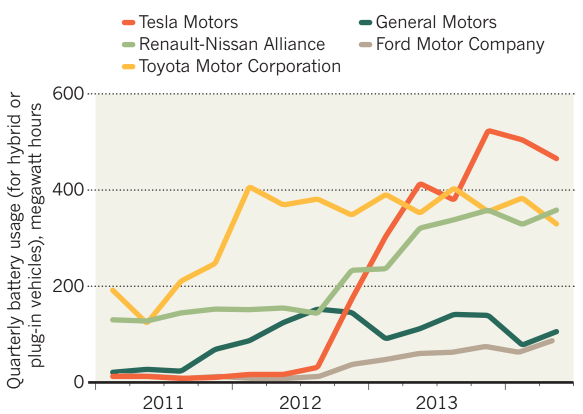
TREND WATCH

SOURCE: LUX RESEARCH

Ending an intense five-state competition, electric-car company Tesla Motors announced on 4 September that it had chosen Reno, Nevada, as the site of a US$5-billion battery 'gigafactory'. Tesla, headquartered in Palo Alto, California, will partner with Japanese battery-maker Panasonic in the endeavour. Tesla, which already dominates the electric-vehicle-battery market (see chart), says that it hopes to cut the costs of battery packs by 30% and to sell 500,000 electric vehicles by 2020.

TESLA'S BATTERY EMPIRE

The California-based electric car maker is one of the largest users of automotive batteries.



NATURE.COM
For daily news updates see:
www.nature.com/news

NEWS IN FOCUS

SPACE Plate tectonics on Jupiter moon buoys bid for trip to Europa **p.153**

CHINA Scientist weighs in on earthquake–reservoir link **p.154**

TECHNOLOGY Astronauts to take delivery of their own 3D printer **p.156**

PHYSIOLOGY Fitness trackers fuel wildlife biology craze **p.157**

CHRIS PANCEWICZ/ALAMY

People in Scotland will be voting on whether they want their country to remain as part of the United Kingdom.

POLITICS

Scientists split over Scottish independence vote

Research could founder or flourish if Scotland leaves the United Kingdom.

BY ELIZABETH GIBNEY

Dolly the cloned sheep was created there; the existence of the Higgs boson was predicted there. But soon Scotland could leave the United Kingdom, with potentially major repercussions for science. Ahead of a historic referendum on 18 September, which the latest opinion polls suggest could go either way, researchers on both sides of the border are split over whether science in Scotland would flourish or founder should its people vote yes to independence.

Although many scientists are reluctant to speak out in a political debate that is fast becoming acrimonious, a few are weighing in. Meanwhile, rival university-based groups, 'Academics for Yes' and 'Academics Together', which draw support from across the humanities and sciences, are busy arguing their cases.

Those in the 'no' camp fear a nation turning in on itself, and the end of a status quo that sees Scotland's scientists produce more papers per head, and receive more citations per paper, ▸

than the UK average (see 'Scottish success').

Other academics argue that a 'yes' vote would give Scotland the freedom to devote more money to science and to organize research to better fit the country's needs. They note that Scotland's science benefited from changes pioneered since the formation of the Scottish government in 1999, a form of independence that devolved certain powers to Scotland, including health and education spending.

A border would hinder the open, liquid exchanges under which science thrives.

One thing is clear: the patchwork of sources from which Scottish institutes currently obtain their funding means that a vote for independence would create complexity. These sources include the European Union, with which Scotland would have to renegotiate membership; bodies specific to Scotland, which may be least affected by independence; and sources that pool money from across the United Kingdom.

The focus of most disagreement is the pan-UK government body Research Councils UK (RCUK), which is responsible for sharing out some £2.9 billion (US$4.7 billion) of tax income collected from England, Scotland, Wales and Northern Ireland. In 2012–13, Scottish institutions received an outsized share of this pot: 10.7% of the total RCUK spending — 13.1% if only university research is taken into account — even though Scotland's population is just 8.4% of the UK total and its tax contribution only 9%. Although small, this difference between what Scotland puts in and takes out might mean a net loss for an independent Scotland, says Omid Omidvar, a social scientist at the University of Coventry in England, who has studied the future of science if independence goes ahead.

The Scottish government, which is spearheading the push for independence, says that it will negotiate a formula for still paying into, and receiving money from, a common research funding system, and will make up any shortfall.

Others say that this is wishful thinking. Earlier this year, the RCUK stated that, "Should there be a vote for independence, the current system could not continue". Keeping Scotland in the RCUK would not be practical, says Hugh Pennington, an emeritus bacteriologist at the University of Aberdeen in Scotland, and a leader of Academics Together, which opposes independence. The desire to ensure that each country gets out what it pays in would cloud decisions and make it difficult to allocate funding in a location-neutral manner, he says. "Scotland is walking away. I think the RCUK would say, 'Fund your own research, Sunshine.'"

Other funding sources would also face changes if Scotland were to leave. Members of the Association of Medical Research Charities invested around £1.1 billion in research in 2011, with 13% of it spent in Scotland. One of the association's wealthiest members, the London-based Wellcome Trust, says that it is unlikely to stop funding Scottish projects entirely if the nation votes 'yes', but that it would review the eligibility of institutions there.

Another issue is ease of access to world-class infrastructure, such as the RCUK-funded Diamond Light Source synchotron near Didcot in England. The resource is used by a range of disciplines for studying matter at the molecular and atomic level. Omidvar says that big questions remain over who would own such sites, which are currently shared between the countries that make up the United Kingdom.

English geneticist and Nobel prizewinner Paul Nurse, who is president of the Royal Society in London, told an audience at the University of Edinburgh in July of his fears that establishing a UK–Scotland border would hinder the "open, liquid and dynamic exchanges" under which science thrives.

Scotland's influence over decision-making in science could be affected, too. Pennington believes that if an independent Scotland joined the European Space Agency, say, the country would become "a relatively small player in a big club, where a few big members have the loudest voices". And some scientists fear that research might become more parochial. One senior Scottish scientist who opposes independence, and who did not want to be named, said that he preferred decisions on research funding to come from London or Swindon, where the RCUK is based, rather than from a small scientific community that would be subject to the sway "of a few very dominant personalities".

Those in the 'yes' camp dismiss such fears, and emphasize the opportunities presented by independence. Bryan MacGregor, a land economist, vice-principal at the University of Aberdeen and a member of Academics for Yes, says that he sees this as a chance to devote a greater portion of government money to research. "In the United Kingdom, we're already spending less on research and development than virtually all our competitors," he says. He notes that further cuts in public spending are planned. "I don't see how the science budget won't be hit."

The trend for innovation would continue under full independence.

Independence would also give Scotland more leverage over science policy and spending, according to the Scottish government. Decisions on taxation and how to allocate tax credits, in which companies receive financial incentives for investing in science, are currently controlled centrally. MacGregor sees a chance to boost the amount that businesses spend on science in Scotland.

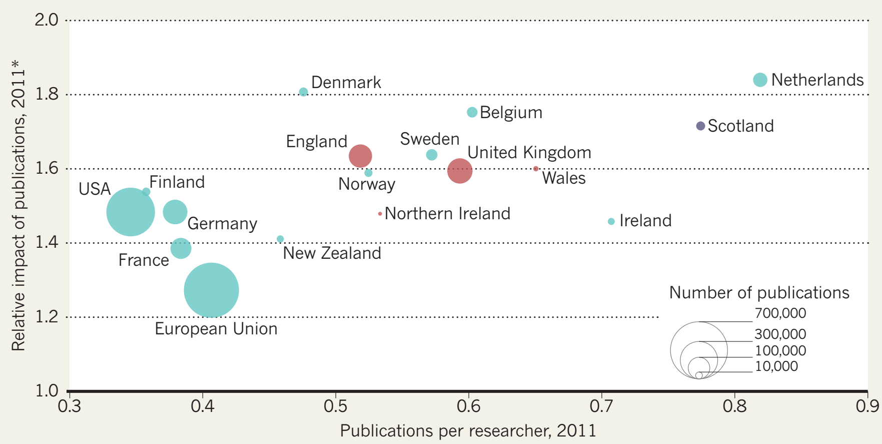
The creation of the Scottish government has already had positive effects on Scottish science. The government has pioneered some original approaches, including the creation of innovation centres, which support collaborations between universities and businesses, and 'research pools' — discipline-specific networks that span different institutions. The trend would continue under full independence, says Murray Pittock, a professor of English who is a vice-principal of the University of Glasgow in Scotland and a member of Academics for Yes.

An independent Scotland would reverse UK immigration policies that the Scottish government says are damaging universities. These include reinstating post-study work visas, which allowed foreign students to work in the United Kingdom for two years once they had finished their degrees and which was scrapped in 2012.

MacGregor says that many of the statements coming from the 'no' campaign are based on fear, and overlook the potential opportunities. But for Pennington, risking a successful research system for the promise of independence is simply not worth it: "If the [Scottish] government had put in its white paper that they were going to raise the university budget threefold, I might have reconsidered my position. At best, what they offer is no change." ■

SCOTTISH SUCCESS

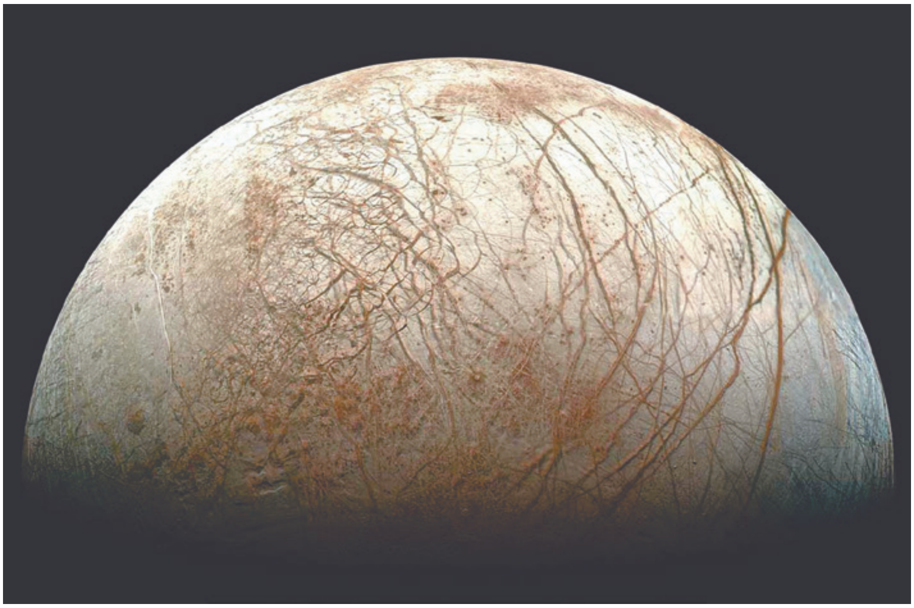
Scottish researchers are more productive than researchers in the United Kingdom as a whole and in many other countries.



*Mean citations per article, normalizing for subject field and year of publication; relative to world average = 1.

SOURCE: ELSEVIER/GO.NATURE.COM/NNLJNJ

GALILEO PROJECT/JPL/NASA; REPROCESSED BY TED STRYK



Some of the scars on the Jovian moon Europa could be the result of subducting plates.

PLANETARY SCIENCE

Plate tectonics found on Europa

Discovery buoys bid for mission to Jovian moon.

BY ALEXANDRA WITZE

If you have got an idea for how to study Europa, then NASA wants to hear from you.

The agency has no official plans for a mission to the Jovian moon, whose icy crust covers a watery ocean in which life could theoretically exist. But spurred by intense congressional interest and several recent discoveries, NASA is seeking ideas for instruments that could fly on a mission to Europa. The possibilities range from a stripped-down probe that would zip past the moon, to a carefully designed Jupiter orbiter that would explore Europa over many years.

The groundswell of enthusiasm is likely to be bolstered by the latest big news, reported on 7 September, that there may be giant plates of ice shuffling around on Europa — much as plates of rock do on Earth (S. A. Kattenhorn and L. M. Prockter *Nature Geosci.* http://dx.doi.org/10.1038/ngeo2245; 2014). Such active geology suggests that Europa's icy surface is connected to its buried ocean — creating a possible pathway for salts, minerals and maybe even microbes to get from the ocean to the surface and back again.

Simon Kattenhorn, a geologist previously at the University of Idaho in Moscow, and Louise Prockter, a planetary scientist at the Johns Hopkins Applied Physics Laboratory in Laurel, Maryland, made the finding after combing through pictures from NASA's Galileo spacecraft, which orbited Jupiter from 1995 to 2003. Most of its pictures of Europa are fairly blurry, but Kattenhorn and Prockter scrutinized one of the few regions of the moon for which high-resolution images exist.

They treated the images as though they were parts of a giant geological jigsaw puzzle, with ridges and bands and other features that have been split and separated by crustal movements, and tried to trace how the surface of Europa had transformed over time. "When we moved all the pieces back together, there was a big hole in the reconstruction, a sort of blank space," says Kattenhorn. The missing portion, the scientists concluded, must have been somehow sucked down into the moon's interior.

Kattenhorn and Prockter propose a system of plate tectonics that involves a shell of ice a few kilometres thick sliding around on warmer, more fluid ice. When one plate hits another and begins to dive downwards — or subduct — it melts and becomes incorporated in the underlying ice, the duo proposes.

Places have already been spotted on Europa where fresh ice crust is being born, but the latest research is the first to pinpoint where it might be going to die.

But without high-resolution images from more areas, researchers cannot tell whether subduction might also be happening in other locations. If it turns out to be common, it might mean that the moon could be cycling life-friendly compounds between the surface and the deep, and that substantially increases the chance that its ocean is habitable, says Michael Bland, a planetary scientist at the US Geological Survey in Flagstaff, Arizona.

The discovery adds to excitement set off in December, when scientists reported plumes of water vapour spurting out at Europa's south pole (L. Roth *et al. Science* **343,** 171–174; 2014). The plumes have not been seen since, and they may or may not be related to Europa's newly appreciated system of plate tectonics. NASA now needs to figure out what kind of mission might best to explore these discoveries.

For the past several years, engineers at the Jet Propulsion Laboratory in Pasadena, California, have been refining a mission concept known as the Europa Clipper. After repeated streamlining, they have come up with a US$2-billion spacecraft design that could carry a range of instruments to the moon (see 'Eye on Europa').

But spooked by the cost, NASA has called for ideas that would run at just $1 billion. The agency is now reportedly evaluating a handful

MORE ONLINE

P. BAUDON/ESA/CNES/ARIANESPACE

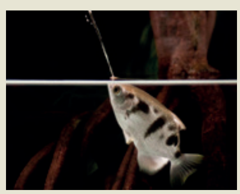
EXPLAINER

Can Europe's GPS be fixed? go.nature.com/tbbggq

MORE NEWS

- Blood transfusions called a priority in fight against Ebola go.nature.com/jgb9hx
- Blue whale numbers making full recovery go.nature.com/oz9klk
- Caffeine evolved at least twice go.nature.com/nfjyyh

VIDEO OF THE WEEK



How archer fish aim their squirt gun go.nature.com/qcvjwd

INGO RISCHAWY/UNIV. BAYREUTH

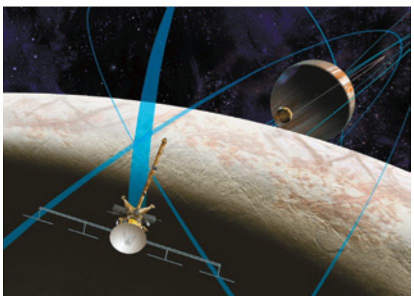
of suggestions.

The strategic down-shift has disappointed some scientists. "It's really frustrating to talk about $1-billion concepts" as if researchers hadn't already considered that, says Britney Schmidt, a planetary scientist at the Georgia Institute of Technology in Atlanta who worked on the Clipper idea. "If you want to do the best science out there, totally committed to by the community, this is the mission you send." Because the Clipper would carry a range of instruments, it could investigate subduction zones, explore plumes and respond to a variety of other research questions, says Prockter.

And in July, NASA asked planetary scientists to submit ideas for instruments they would like to see fly onboard a craft such as the Clipper, whatever the cost. Proposals are due by 17 October, and the agency plans to choose around 20 of them next April for further development.

Although NASA is worried about the overall cost of a Europa mission, it has money to spend in the short term. For each of the past couple of years, Congress has given the agency's planetary-sciences division tens of millions of dollars more than it asked for, and directed it to spend the money on Europa mission concepts. The drive is led by Congressman John Culberson (Republican, Texas), a *Star Trek*-quoting space enthusiast who sits on a powerful spending committee.

Europa researchers are happy to take advantage. "I'm frothing at the mouth in excitement," says Kattenhorn. "There is clearly so much more that we still need to learn about Europa." ■



NASA/JPL-CALTECH

EYE ON EUROPA

A mission being considered by NASA would carry a range of instruments to explore a variety of questions at the Jovian moon Europa.

RADAR	Characterize the moon's icy shell
SPECTROMETERS	Study surface chemistry
CAMERAS	Image the surface and its topography
MAGNETOMETER	Probe the moon's interior
LANGMUIR PROBE	Explore environmental plasmas
THERMAL INSTRUMENT	Investigate heat flow

SUN RONGGANG XINHUA NEWS AGENCY/NEWSCOM

The Xiluodu hydropower station is sited directly over one of China's many fault lines.

HYDROPOWER

Chinese data hint at trigger for fatal quake

Seismic activity started to rise just as two giant reservoirs on upper Yangtze were being filled with water.

BY JANE QIU

Ever since 3 August, when an earthquake in southwestern China killed more than 600 people, the Chinese media and blogosphere have buzzed with speculation that the magnitude-6.5 tremor was linked to the filling of two gigantic reservoirs along the upper Yangtze river. Now, a geologist says that he has data to support the possible link.

On 28 August, Fan Xiao, an engineer at the Sichuan Bureau of Geology and Mineral Resources in Chengdu, reported a rough correlation between the timing of the filling of the reservoirs and a rise in seismic activity in the surrounding region.

Posted on a website run by Probe International, a non-profit organization that reports on China's large-scale water projects, Fan's analysis is based on crude seismic data — the only sort that are publicly available — so the link is tentative. But "it's an important possibility", says Hu Xian-ming, a geophysicist at the Sichuan Earthquake Administration in Chengdu. "There are serious concerns for deadly quakes in the future."

Criss-crossed by active faults, the upper Yangtze region is seeing a boom in dam-building for the generation of hydropower. But when water flows quickly into the resulting reservoirs, it can change the stress on faults deep underground, either from the sheer weight of the water, or when water infiltrates the rocks through cracks and pores. These events might accelerate a fault's natural 'seismic clock', hastening an earthquake that is already building, or increase the chance of one occurring at all.

Debate is already raging about whether the 2008 quake in Wenchuan county, which killed at least 70,000 people (see *Nature* **459,** 153–157; 2009) was linked to the filling of Zipingpu reservoir in Sichuan province. Fan was one of the first to raise the possibility, and his suggestion was followed up by other researchers who have reported, for instance, that the reservoir might have brought forward the occurrence of the quake by tens to hundreds of years (S. Ge *et al. Geophys. Res. Lett.* **36,** L20315; 2009).

After the 3 August quake in Ludian county, discussion turned to two newly created

reservoirs, the nearest of which, Xiluodu, is 40 kilometres from the epicentre (see 'Of dams and quakes').

On the basis of seismic readings taken between January 2010 and July 2014, Fan reports that small quakes became more frequent in late 2012 and continued until the end of the period. That heightened activity roughly correlates with the reservoirs being filled. The areas most affected cluster at three locations: one near each reservoir, and a third close to a fault whose rupture led to the latest quake. "The study has its limitations," says Fan, "but it does ring an alarm bell of increasing reservoir-triggered quakes in the region."

OF DAMS AND QUAKES

Since a magnitude-6.5 tremor hit Ludian on 3 August, speculation has been rife about whether filling of the Xiluodu and Xiangjiaba dam reservoirs played a part.

A POSSIBLE TREND

His report also flags up two smaller earthquakes that hit the county of Yongshan in April and August and were caused by faults directly below the Xiluodu reservoir. Xu Xi-wei, deputy director of the China Earthquake Administration's Institute of Geology in Beijing, agrees that these two tremors were "most likely triggered by Xiluodu". The link between the Ludian quake and the reservoirs is less convincing, he says, because the epicentre is too distant and the initial rupture happened about 12 kilometres down, too deep for water to reach.

However, Christian Klose, a geologist at Think Geohazards, a consultancy firm in Bronxville, New York, who studied the Wenchuan quake, says that he finds the link to the Ludian quake plausible. "You don't need water migrating into rocks to cause quakes," he says. "The sheer weight of a massive reservoir could bend Earth's crust and rupture a critically stressed fault."

Researchers including Hu are calling for the dissemination of more-sensitive seismic data from the dense network of stations in the reservoir areas, which are tightly controlled by hydropower companies. "This would make more detailed analyses possible," he says. Fan's analysis was based on data from seismic stations controlled by the regional government, which are sparse and have less-sensitive equipment.

With dozens more dams planned or under construction for the upper Yangtze, "the issue is more pressing than ever", says Fan. Whether or not the Ludian quake was triggered by reservoir-filling, it would be prudent to prepare, says Xu. "Buildings in reservoir areas must be reinforced to fend off future quakes." ■

Engineers test a 3D printer under microgravity conditions aboard a modified aircraft in parabolic flight.

MADE IN SPACE

TECHNOLOGY

NASA to send 3D printer into space

Machine will let astronauts create parts to order.

BY ALEXANDRA WITZE

In one small step towards space manufacturing, NASA is sending a 3D printer to the International Space Station. Astronauts will be able to make plastic objects of almost any shape they like inside a box about the size of a microwave oven — enabling them to print new parts to replace broken ones, and perhaps even to invent useful tools.

The launch, slated for around 19 September, will be the first time that a 3D printer flies in space. The agency has already embraced ground-based 3D printing as a fast, cheap way to make spacecraft parts, including rocket engine components that are being tested for its next generation of heavy-lift launch vehicles. NASA hopes that the new capability will allow future explorers to make spacecraft parts literally on the fly.

Space experts say that the promise of 3D printing is real, but a long way from the hype that surrounds it. "There's been a tendency among the space-enthusiast crowd to treat 3D printing as if it's a magic technology — as if all you have to do is wave your wand, say 'Abracadabra, here's a 3D printer', and it's going to build you a Moon base," says Dwayne Day, a senior programme officer at the National Research Council in Washington DC who oversaw a recent report on 3D printing in space (see go.nature.com/j6z5mq). In reality, Day says, the technique "is an important component of a much broader technology base that is being developed and advanced".

The printer selected by NASA was built by the company Made in Space, which is based at a technology park next to NASA's Ames Research Center in Moffett Field, California. During the printer's sojourn on the space station, it will create objects from a heat-sensitive plastic that can be shaped when it reaches temperatures of about 225–250 °C. The team is keeping quiet about what type of object it plans to print first, but the general idea is to fashion tools for use aboard the station. "If you have 300 different things that could break on your spacecraft, you may not need to carry replacement parts for all 300 of them," says Day.

The Made in Space printer is also a testbed for performance of the technology in near-zero gravity. The machines work by spraying individual layers of a material that build up to form a complete, 3D object. But in near-weightless environments, there is no gravitational pull to hold the material down.

In test flights aboard 'vomit comet' aircraft that fly in a parabola to create almost weightless conditions, Made in Space discovered that the layers of printed material varied substantially in thickness as the aeroplane cycled in and out of microgravity. By modifying the printer, the team got the layers to come out at roughly the same thickness.

Thermal issues could also be a problem. Heat flows differently in microgravity, which could mean that parts of the plastic become too hot or too cool for the printing to work properly. "Whether it works fantastically or we have some issues, we're going to learn things that will play into the design of future machines," says Michael Snyder, the company's director of research and development.

Made in Space is looking at flying a second printer to the space station next year, incorporating design changes from what is learned during the first flight. To evaluate the printer's performance, parts made aboard the space station this time will be flown back to Earth and tested to see whether they work as well as Earth-made materials do. There is little point in manufacturing parts in space if they do not work at least as well as spares that an astronaut might grab from a storage locker, Day notes. Time is also an issue: Made in Space's prints typically take between 20 minutes and two hours, which might not be useful, depending on the urgency of the situation.

Back on Earth, NASA has already found ways in which 3D printing (also known as additive manufacturing) might save time and money. At NASA's Marshall Space Flight Center in Huntsville, Alabama, engineers have been testing 3D-printed components for rocket engines that use liquid propellant. In one recent project, metal printouts of a rocket-engine injector cut at least 80% off the cost of the US$300,000 part. Another upcoming test will evaluate a 3D-printout of an even more complex piece of machinery: the fuel turbopump that serves as the heart of a rocket engine. "We're looking to apply additive manufacturing where it makes sense," says Nick Case, an engineer at Marshall.

And designers can use 3D printing to produce shapes that have never been seen before in spacecraft, says Slade Gardner, an engineer at Lockheed Martin Space Systems in Littleton, Colorado. "We have a long way to go, and we can't do everything with 3D printing," he says. "But the real, long-term goal is a design revolution." ■ SEE EDITORIAL P.144

CORRECTION

The graphic in the News story 'Ebola drug trials set to begin amid crisis' (*Nature* **513,** 13–14; 2014) said that the NIAID/GSK vaccine had been shown to be safe in 'preclinical human trials'. In fact, two components of the vaccine have been shown to be safe in other phase I trials, and the vaccine itself is about to enter phase I clinical trials.

TERRIE M. WILLIAMS (NATL MARINE FISHERIES SERVICE PERMIT #13602-01)

SURVIVAL OF THE FITTEST

Using a wildlife version of fitness trackers, biologists can finally measure how much energy animals need to stay alive.

BY ANDREW CURRY

After attacking more than a dozen young Hawaiian monk seals, the aggressive male known as KE18 had to be banished from his island home. He now lives in a seawater pool in California, a few hundred yards from the Pacific Ocean.

Instead of assaulting more members of his species — one of the most endangered marine mammals on the planet — KE18 is now providing information that could help to keep the seals from going extinct. On a breezy June day at the University of California, Santa Cruz, an animal trainer is working to teach him to wear a metal tube containing a trio of accelerometers, tiny devices that track and log changes in velocity. It is the seal version of the fitness trackers that joggers and other endorphin addicts wear on their wrists to capture their every move.

Heaving his 200 kilograms up on the side of the pool, KE18 opens his mouth to show off his teeth and gets a fish as a reward. "We use a lot of fish and hot dogs in our lab," says biologist Terrie Williams, who heads the mammalian physiology lab.

Seal KE18 is trained in Santa Cruz, California, to swim with a fitness tracker on a flipper.

The hope is that data collected on KE18 will help to explain a mystery about the threatened Hawaiian monk seal (*Monachus schauinslandi*). Most of the population lives in the remote northwestern end of Hawaii's island chain, where four in every five pups die before reaching adulthood and the population is falling by more than 3% a year. "They end up starving to death," says Charles Littnan, who runs the Hawaiian Monk Seal Research Program at the National Oceanic and Atmospheric Administration in Honolulu. But a much smaller population living close to the crowded beaches of the main Hawaiian islands is actually growing, despite all the human activity nearby.

KE18 could be the key to deciphering why the two groups are on ▶

opposite trajectories. Researchers at the university's marine lab are teaching the seal to wear a flipper cuff holding the accelerometers while exercising in his pool. Soon, they will also start to measure the oxygen content of his breath — a proxy for how hard he is working. By combining those data and the accelerometer information, Williams and her colleagues can calibrate the amount of energy that KE18 uses as he swims. They plan to use those data to interpret the readings from accelerometers placed on wild monk seals, which should help to determine why they are not thriving — whether they are wasting energy, for example, by having to swim too far to catch prey.

Seals are not the only animals wearing fitness trackers. In the Arctic earlier this year, researchers fitted polar bears (*Ursus maritimus*) with accelerometers to see how much energy they expend while swimming the ever-increasing distances to reach dwindling ice floes. Marine biologists are deploying the devices to find out whether warming oceans are impairing the swimming prowess of fish. And in the western United States, motion trackers on mountain lions (*Puma concolor*) will help to determine how much extra energy they burn negotiating the sprawl of human development.

As accelerometers have become cheaper and smaller — largely because of their use in mobile phones — wildlife biologists have embraced them as a way to collect data on movement and energy consumption. The information is starting to answer basic questions about animal behaviour and physiology, and help researchers to predict how climate change, habitat destruction and human activity will affect animals.

The flood of data coming in is altering the kinds of questions that wildlife biologists ask, says Craig Franklin, an eco-physiologist at the University of Queensland, Australia. "People are really starting to realize the value of physiology in addressing conservation."

"IF ANIMALS ARE GOING TO COLLECT ONE THING THAT'S ANALOGOUS TO MONEY, IT'S ENERGY."

THE COST OF ENERGY

Biologists sometimes say that energy is the currency of life. "If animals are going to collect one thing that's analogous to money, it's energy," says Rory Wilson, an aquatic biologist at Swansea University, UK. Everything that an organism does requires some expenditure. Movement is particularly costly; but even asleep, the body burns calories to maintain digestion, breathing and circulation.

There is a key difference, however, between money and energy: animals cannot go into the red. "If you don't have energy in your system, you're dead," says Wilson.

Biologists do not have an easy way to track how much energy wild animals use, so they use oxygen consumption as a proxy. The more fuel that animals burn, the more oxygen they need.

With an animal on a treadmill, researchers can correlate oxygen use with the speed of the machine to come up with a rough measure of energy use per metre of forward motion, for example. That 'cost of transport' has long served as a crude yardstick of animal energetics. But it relies on many dubious assumptions, such as that animals move through the environment much as they run on a treadmill — when in fact they could burn much more when covering uneven ground. The technique is even less useful for animals that crawl through sand, swim or fly. And because it relies on knowing the distance travelled, it is hard to apply to animals that spend a lot of time obscured underwater or in the dark, for example.

That's where accelerometers come in. First widely applied as the sensors that trip air bags in cars, accelerometers contain tiny weights that shift with a change in speed. When biologists initially packaged them with data loggers and batteries to measure animal motion, the devices were unwieldy. In the late 1990s, Williams put flipper cuffs and the equivalent of small scuba tanks on Weddell seals (*Leptonychotes weddelli*) in Antarctica[1]. "It was a pretty exotic piece of equipment," she says.

➲ NATURE.COM
See videos and a slideshow of fitness tracking in animals:
go.nature.com/3cdpf8

As costs plummeted, however, companies began to offer ready-made accelerometers combined with memory chips that record dozens of data points each second for weeks or even months at a time. And wildlife biologists immediately saw an opportunity to fix them to animals. Since 2009, more than 130 papers have been published using accelerometers to study animal behaviour. "It's really exploded in terms of interest and technology," Williams says. "Now we can sample so much, we know every time an animal takes a stroke or a paw hits the ground."

Such detailed results led Wilson to wonder whether accelerometers might be the key to a more flexible and accurate measure of energy use. In 2005, he worked with Lewis Halsey, an environmental physiologist now at the University of Roehampton, UK, to put accelerometers on a quintet of great cormorants (*Phalacrocorax carbo*). They coaxed the diving birds to walk on a treadmill inside a sealed respirometry chamber while wearing a 35-gram data logger the size of a book of matches.

The resulting data confirmed that the set-up worked as a way to measure energy expenditure: when the accelerometers recorded more movement, oxygen consumption rose accordingly[2]. The motion information was useful outside the lab, too. Accelerometers taped to wild cormorants revealed that birds carrying a load of fish required 14% more acceleration to stay aloft than did unladen birds, for example.

Scientists now refer to dynamic body acceleration, or DBA, a tally of acceleration in three dimensions. Since Wilson and Halsey's experiments with cormorants, there has been a flood of publications using DBA and oxygen consumption to investigate the energy demands of wild animals from lobsters to badgers, toads and penguins. Other labs are working on commercially important species such as scallops, cod and sea bass. "Everyone's saying 'bloody hell, it works on everything'," says Wilson.

To test the effects of climate change on fish, behavioural ecologist Julian Metcalfe of the Centre for Environment, Fisheries and Aquaculture Science (CEFAS) in Weymouth, UK, implanted accelerometers into cod (*Gadus morhua*) and studied how the fish responded metabolically to different temperatures. "If the water gets colder, they get less active — but are still able to chase and escape when necessary," he says.

When the temperature rises, however, cod have trouble performing high-energy feats. "A fish at high temperatures is less able to escape a predator or catch prey," Metcalfe says. "In terms of climate change, that's important to know."

Serena Wright, now a biologist at CEFAS in Lowestoft, UK, studied farmed trout with the help of implanted accelerometers. She found that some of the fish had stunted fins or deformed skeletons, perhaps because of crowding or inbreeding, and the deformed fish used more energy than normal fish while swimming, which slows their growth. The fish had also learned to anticipate regular feedings: they moved relatively little until just before meal times, she says. That is a problem because aquaculture fish grow more quickly when they swim. Wright has therefore suggested that fish farms adopt irregular feeding schedules to keep the fish moving more during the day.

BREEDING DISTURBANCES

Anthony Robson, a biologist at the University of Western Brittany in Brest, France, and Halsey have even attached motion sensors to largely sedentary creatures, such as scallops (*Pecten maximus*). The shellfish move only about two minutes per day, on average — but that accounts for around 17% of their energy expenditure in the wild[3]. Farmed scallops, by contrast, spend more time in escape and evasion mode than their wild brethren because of the lights and vibrations from human activity, and that bumps up their energy expenditure to around 40%. "They're basically going crazy being disturbed in the hatchery," says Robson. Rather than feeding and growing, the farmed scallops spend much of their energy


Movement-tracking devices help researchers to gauge how much energy polar bears burn, and show that human activity impedes the growth of farmed scallops.

MICHAEL DURHAM, COURTESY OF OREGON ZOO/USGS

ERWAN AMICE/CNRS

budget on moving — information that could be valuable for fisheries.

Energy data can also help to explain some of the behaviour that animals exhibit in the wild. Take the Magellanic penguin (*Spheniscus magellanicus*), found in the cold waters off the coast of southern South America. Researchers know that the birds prefer anchovies over squid, which they eat only when the fish are hard to come by. But the reasons were not clear — until Wilson put accelerometers on the birds in the lab and then in the wild. He found that the natural buoyancy of penguins made swimming downwards much more costly than going up[4]. The penguins use less energy catching fish because they can take advantage of their buoyancy and lunge at the prey from underneath. "But penguins have to swim actively after diving squid," Wilson says. "The energy costs for catching a squid are hugely higher than for an anchovy. We never would have got that without accelerometry."

And by looking for tell-tale wiggles in the accelerometer data, the researchers could even tell how many times penguins were feeding in the wild, allowing them to estimate the penguins' total consumption — data that could become important if countries want to limit fish harvests to protect key penguin populations.

In Williams's lab, researchers are going after bigger beasts. PhD student Anthony Pagano, a wildlife biologist with the US Geological Survey, is trying to figure out how to get polar bears into a respirometry chamber to measure their energy use. Those data would complement the measurements he has made over the past two years, as part of a project that put tracking collars on female polar bears in Alaska.

The location readings show that as sea ice retreats, bears are swimming two to three times farther than they did 10 or 20 years ago, says Pagano. In 2008, for example, researchers recorded a polar bear making a 687-kilometre continuous swim over 9 days to reach pack ice. But what researchers do not yet know is how much more energy the bears are using in the process. So Pagano has fitted the animals with collars that will gather acceleration, global-positioning and even video data over the eight-month-long Arctic winter.

Meanwhile, he is working with zoos to put captive females, which weigh up to 370 kilograms, into respirometry chambers built around a polar-bear-sized 'endless pool' or treadmill. "It's going to come down to the ability to build something that's polar-bear proof," he says. "One of the challenges is that they're pretty destructive and big."

Pagano's struggles point out a key obstacle to the burgeoning field's growth: to get accurate measurements of energy expenditure, biologists need to find ways to calibrate the devices on individual species. "It's a hell of a job to calibrate with a new animal," Halsey says.

Wild animals must be trained or coaxed into a respirometry chamber, whether in the field or in a lab, which sometimes proves difficult. Often, the animals simply do not take to treadmills. During her PhD, Astrid Willener found herself wrestling waist-high king penguins (*Aptenodytes patagonicus*) onto a treadmill at a research station on France's Possession Island to study their energetics. Some were too stressed to walk on the treadmill and tried to slide on their bellies or skidded along on the soles of their feet with their backs resting against the wall of the box. "Once a penguin's found a trick, they'll continue," Willener says.

Even fitting the accelerometer takes creative thinking. Williams's team has come up with 'wetsuits' containing accelerometers and heart-rate sensors that they can slip onto captive dolphins, which they teach to surface under breathing domes like the one used to see how much oxygen KE18 uses. And Halsey has used bits of a pair of tights to strap the devices onto cane toads, which puff up when they are handled.

Halsey and Wilson, the pioneers of animal accelerometers, hold out hope that researchers will eventually come up with an algorithm that would allow them to use the data without requiring every species to go into a respirometer. But colleagues are sceptical. "My feeling is that there's so much variation between individual animals that you have to do at least animals within that species," says eco-physiologist Michael Scantlebury of Queen's University Belfast, UK. Pagano's polar bears, for example, are such good swimmers that he thinks no other type of bear would work as a proxy.

Caleb Bryce, another of Williams's graduate students, wants to study the energy budgets of grey wolves (*Canis lupus*). He initially hoped that large dogs would be acceptable substitutes and has put dozens of them on a treadmill inside a sealed transparent box. But because wolves eat much less frequently than domestic dogs and seem to have a different metabolism, Bryce has concluded that he will need to train captive wolves to run on a treadmill to get exact numbers. It is not so far-fetched: Williams and a collaborator in Colorado managed to convince a captive mountain lion to exercise in a respirometry chamber as part of an effort to assess how the animals are faring in the Santa Cruz mountains as humans encroach on their habitat.

Proponents of wildlife accelerometry admit that it will take some time to work out the logistics of how to collect energy data — and how to make use of them. But with animals such as KE18 providing potentially important information that might help to save a species, researchers hold out a lot of hope. "We're just at the beginning," says Williams. "It's a really exciting time for wildlife biology." ■

Andrew Curry *is a freelance writer in Berlin.*

1. Davis, R. W. *et al. Science* **283,** 993–996 (1999).
2. Wilson, R. P. *et al. J. Anim. Ecol.* **75,** 1081–1090 (2006).
3. Robson, A. A., Chauvaud, L., Wilson, R. P. & Halsey, L. G. *J. R. Soc. Interface* **9,** 1486–1498 (2012).
4. Sala, J. E., Wilson, R. P. & Quintana, F. *PLoS ONE* **7,** e51487 (2012).

VISION QUEST

TECHNOLOGIES ARE ALLOWING DOCTORS TO DO WHAT WAS ONCE UNHEARD OF: RESTORE BLIND PEOPLE'S SIGHT. NOW THE REAL CHALLENGES BEGIN.

BY CORIE LOK

Tami Morehouse's vision was not great as a child, but as a teenager she noticed it slipping even further. The words she was trying to read began disappearing into the page and eventually everything faded to a dull, grey haze. The culprit was a form of Leber's congenital amaurosis (LCA), a group of genetic disorders in which light-sensing cells in the retina die off, usually resulting in total blindness by the time people reach their thirties or forties. But Morehouse got a reprieve. In 2009, at the age of 44, the social worker from Ashtabula, Ohio, became the oldest participant in a ground-breaking clinical trial to test a gene therapy for LCA. Now, she says, she can see her children's eyes, and the colours of the sunset seem brighter than before.

Morehouse calls these improvements life-changing, but they are minor compared with the changes in some of the younger trial participants. Corey Haas was eight years old when he was treated in 2008 — the youngest person to receive the therapy. He went from using a white cane to riding a bicycle and playing softball. Morehouse often wonders what she would be able to see now if she had been closer to Haas's age when she had the therapy. "I was born a little too soon," she says.

Visual impairment affects some 285 million people worldwide, about 39 million of whom are considered blind, according to a 2010 estimate from the World Health Organization. Roughly 80% of visual impairment is preventable or curable, including operable conditions such as cataracts that account for much of the blindness in the developing world. But retinal-degeneration disorders — including age-related macular degeneration, the leading cause of blindness in the developed world — have no cure.

SMALL STEPS

In the past seven years, there has been mounting hope and excitement about the prospect of slowing or even reversing vision loss from retinal disorders. Clinical trials testing gene therapy, cell transplants and retinal prostheses are under way, and many studies — including the trial[1,2] involving Morehouse and Haas — are producing promising results. Biotechnology firms are taking up the challenge, and several have formed to take treatments through clinical testing. But most of the successes so far have been in treating rare congenital disorders, and it is still unclear how many people will ultimately benefit and to what extent vision can be preserved or restored. "There's a growing appreciation of the complexity of the clinical problem," says Thomas Reh, a neurobiologist working on cell transplants for the eye at the University of Washington in Seattle.

It may seem vulnerable and complex, but the eye has features that make it a good testing ground for experimental treatments. Unlike internal organs, surgeons can easily operate on it and peer inside to track how well a therapy is doing. It is also walled off from many damaging inflammatory responses that might derail a cell-transplant or a gene therapy. So the eye is "a good way to dip the toes in the water", says Stephen Rose, chief research officer at the Foundation Fighting Blindness in Columbia, Maryland, which funds research and consults with drug firms.

Since 2007, clinical researchers have been dipping their toes into gene therapy for congenital forms of retinal degeneration such as LCA. The goal is to use a virus to supply retinal cells with working copies of a gene called RPE65, which is mutated in the form of the disease known as LCA2. The hope is that the working gene will repair malfunctioning cells and keep them alive, preserving and even improving vision. Trials by three different groups[1–4] have shown not only that the procedure is safe, but also that it boosts vision in most participants — and that most improvements seem to be maintained for up to seven years. The biotechnology company Spark Therapeutics in Philadelphia, Pennsylvania, is now testing gene therapy for LCA2 in an advanced trial, and it hopes to file for regulatory approval in the United States as early as 2016.

> "WE'RE STILL AT THE VERY START OF THE OPTIMIZATION PROCESS IN HUMANS."

But some studies have raised questions about how well the therapy is working. An analysis[5] of data from one[4] of the three initial trials found that although participants' vision had improved, their photoreceptors were still dying at about the same rate as before the treatment. Artur Cideciyan, a vision scientist at the University of Pennsylvania in Philadelphia and a co-author of the study, says that the improvements probably came from the rescue of only some retinal cells. Gene therapy may not have affected the more dysfunctional photoreceptors, and these were the ones that probably died after treatment.

Researchers have observed that there seems to be a point of no return in some forms of retinal loss[6]. A possible reason is that cell death disrupts the structure of retinal tissue, leading to a domino-like decline. Cideciyan argues that after retinal degeneration has started, even cells improved by gene therapy may eventually die off, at least in LCA2.

Robin Ali, a geneticist at University College London who led one of the early LCA2 trials[3], is more confident about the promise of these treatments: the careful animal work that preceded human trials showed that gene therapy, if given at the right dose and the right time, can slow degeneration. "We're still at the very start of the optimization process in humans," he says.

Discovering the best timing for the treatment in humans remains a central challenge. Most researchers agree that the best approach is to replace the faulty gene when patients are young, before the degeneration starts or at least when there are more viable cells to save. That could mean doing retinal surgery in someone with good vision — a difficult decision, says Robert MacLaren, an ophthalmologist at the University of Oxford, UK, who is running a gene-therapy clinical trial for another form of congenital blindness[7]. "That is where the risks are greatest, but so are the gains." Spark's phase III LCA2 trial was open to children as young as three years old.

Once the damage has reached a point at which there are few or no useful photoreceptors left to save, gene therapy will probably not help. That is why some research groups are looking to other techniques, such as cell-based therapy.

REGENERATION GAME

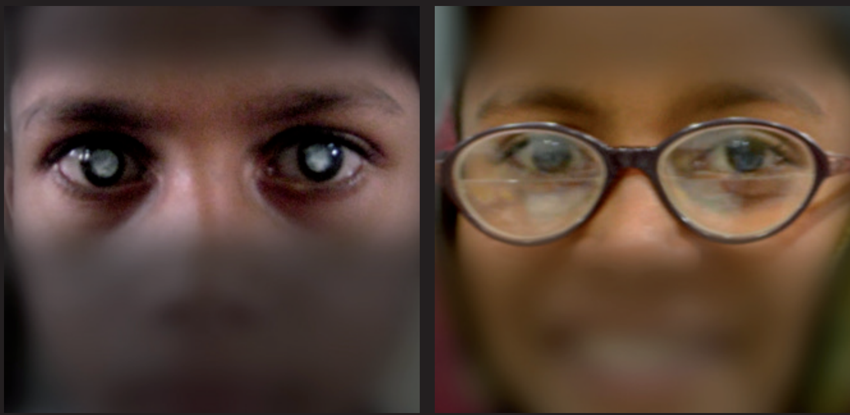
When people talk about the therapeutic potential of embryonic stem cells, they usually mention treatments for diabetes and spinal-cord injuries. But one of the first clinical trials for such cells was actually to treat blindness. Advanced Cell Technology in Marlborough, Massachusetts, has been conducting trials that transplant retinal pigmented epithelial (RPE) cells derived from embryonic stem cells into people with one of two forms of vision loss caused by retinal degeneration (see *Nature* **481,** 130–133; 2012). The trials started in 2011, and researchers and industry onlookers are eagerly anticipating results later this year.

The RPE cells support the function of the photoreceptors, and the hope has been that the cell transplants will stop or slow the loss of the light sensors. Replacing photoreceptors themselves could have a higher pay-off, but deriving them efficiently from stem cells and wiring them into the retina has been difficult.

There are tantalizing signs that it could work. Ali and his colleagues, for example, have shown that when precursors to rod cells — photoreceptors that are active in dim light — are transplanted into mouse eyes, they connect with other cells in the retina and can restore vision[8]. They also showed that the rods can be grown from mouse embryonic stem cells and can mature and integrate into the retina[9]. Researchers are now working on deriving and transplanting cone cells — which enable high-acuity vision — into animals, and are starting to think about the first human trials.

The Argus II artificial retina allows patients to distinguish light from dark, but does not yet restore full vision.

PATRICE LATRON/LOOK AT SCIENCES/SPL



PAWAN SINHA

Pre- and post-operative photos of a child born blind owing to dense cataracts.

A critical question

In the 1960s, neuroscientists showed that if one of a cat's eyes is sutured shut early in life, then the animal will always be blind in that eye[10]. This led to the idea of a 'critical period' in visual development, a time during which visual circuitry must be used, or it will never work properly (see *Nature* **487,** 24–26; 2012). Now, however, people who have had their vision restored are providing neuroscientists with a fresh opportunity to examine the critical period and monitor how the brain responds to visual signals that it has long been deprived of.

Leading the way is Project Prakash, an organization that has delivered vision care to more than 1,400 children in rural India since 2003. The project, headed by neuroscientist Pawan Sinha at the Massachusetts Institute of Technology in Cambridge, has given sight to more than 450 children who were born blind because of cataracts, but underwent operations to remove them when they were children or teenagers — long after the critical period for visual development was thought to have passed. Sinha and his colleagues have found[11] that some aspects of the children's vision — such as visual acuity, which is needed for reading — seem to be permanently impaired, but that others, such as the ability to tell a face from a non-face, show some level of recovery.

This shows that the critical period is not absolute, and that a person's brain can develop significant vision even if it is first exposed to visual signals relatively late in life. "It's not the case that they are completely compromised," says Sinha.

Another study[12] has shown how the human visual system remains resilient later in life as long as damage to the retina can be repaired. A group led by Manzar Ashtari, a brain-imaging specialist at the Children's Hospital of Philadelphia in Pennsylvania, carried out brain-imaging studies on people whose vision had been partly restored during a gene-therapy clinical trial for a congenital form of retinal degeneration. They found that even after up to 35 years of severely impaired vision, the study participants were, surprisingly, still able to use the neural circuitry that is normally used for vision, says Ashtari. "The pathway is still intact after years of deprivation," she says.

Recipients of treatment may require some training and therapy, but these studies are cause for optimism. If the eye can be fixed, the brain's visual system could be malleable enough to turn light signals into useful sight. **C.L.**

Whatever the strategy, stem-cell-based treatments face some of the same issues as gene therapy: the disease processes that kill off retinal cells could continue to do so after treatment. There may be ways around this for less severe forms of blindness, says Ali, but transplanting cells into the eyes of people with very advanced disease might not work. So for these people, a more radical solution may be required.

When doctors first turned on his bionic eye, Roger Pontz thought he was dreaming: for the first time in 15 years, he could see the lights on the ceiling. The 56-year-old dishwasher from Reed City, Michigan, is one of 90 people worldwide to have received the Argus II implant, the only approved retinal prosthesis on the market. Pontz lost his sight to retinitis pigmentosa, a group of inherited disorders that cause retinal-cell death and leave most patients legally blind by the age of 40. Now, he no longer bumps into walls and he can grab the refrigerator door handle without having to feel his way towards it. "It's made life a lot better," says Pontz.

The Argus II, made by Second Sight in Sylmar, California, was approved by the US Food and Drug Administration in 2013 for severe retinitis pigmentosa. It consists of a small camera mounted on a pair of glasses, which sends video data to a portable computer worn by the user. The processed signals are sent back through a cable to the glasses, where they are then transmitted wirelessly to a receiver wrapped around the eye. This sends the signals to a chip that has been surgically placed on top of the retina. The chip generates electrical impulses that stimulate the remaining cells of the retina.

The device — which is pricey at US$144,000 — does not restore normal vision. "We try to transform blind people into low-vision people," says José-Alain Sahel, an ophthalmologist and head of the Institute of Vision in Paris, who was involved in testing the device in humans. Pontz says that he sees dots of light in black and white, which correspond to lines of contrast such as a doorway. With rehabilitation exercises, he is learning to make sense of those patterns (see 'A critical question'). He still uses a white cane and has to move his head continually up and down and side to side so that the camera in his glasses can take in the scenery.

Second Sight is now aiming to open up the technology to more people. The firm hopes to begin testing the Argus II in people with age-related macular degeneration this year. To boost the device's resolution, the company has tried squeezing more electrodes onto the chip but that did not make much of a difference. So, instead, it is focusing on improving the software, with some promising early results.

With so many advances, researchers are optimistic about the future. Even if a treatment can save or restore only a small number of light sensors in a diseased retina, that might still be enough, says Ali. "You don't need very many functioning photoreceptor cells for vision."

It is probably not going to be perfect vision, and it may not even be a permanent fix, but as recipients such as Morehouse say, every little bit of improvement is significant.

"Even if you can give me five to ten years of vision, I'm going to take it." ■

Corie Lok *is* Nature*'s Research Highlights editor.*

1. Maguire, A. M. *et al. Lancet* **374,** 1597–1605 (2009).
2. Maguire, A. M. *et al. N. Engl. J. Med.* **358,** 2240–2248 (2008).
3. Bainbridge, J. W. B. *et al. N. Engl. J. Med.* **358,** 2231–2239 (2008).
4. Hauswirth, W. W. *et al. Hum. Gene Ther.* **19,** 979–990 (2008).
5. Cideciyan, A. V. *et al. Proc. Natl Acad. Sci. USA* **110,** E517–E525 (2013).
6. Cepko, C. L. & Vandenberghe, L. H. *Hum. Gene Ther.* **24,** 242–244 (2013).
7. MacLaren, R. E. *et al. Lancet* **383,** 1129–1137 (2014).
8. Pearson, R. A. *et al. Nature* **485,** 99–103 (2012).
9. Gonzalez-Cordero, A. *et al. Nature Biotechnol.* **31,** 741–747 (2013).
10. Hubel, D. H. & Wiesel, T. N. *Naunyn-Schmiedebergs Arch. Exp. Pathol. Pharmakol.* **248,** 492–497 (1964).
11. Sinha, P. & Held, R. *F1000 Med. Rep.* **4,** 17 (2012).
12. Ashtari, M. *et al. J. Clin. Invest.* **121,** 2160-2168 (2011).

COMMENT

CLIMATE China reframes emissions debate by leading renewables revolution **p.166**

HISTORY Mind–body dualism haunts French science fiction **p.169**

SCI-FI Neal Stephenson on an anthology to stimulate engineers **p.170**

EPIGENETICS Study the effect of pregnancy on mothers' health, not just babies' **p.172**

ILLUSTRATION BY DAVID PARKINS

Rethink IPCC reports

Voluntary work alone cannot sustain the assessments carried out by the Intergovernmental Panel on Climate Change. **Thomas F. Stocker** and **Gian-Kasper Plattner** call for institutional support and a longer report cycle.

Working on an assessment for the Intergovernmental Panel on Climate Change (IPCC) is utterly exhausting. Most authors are proud of their team's achievement and enjoy the intense discussions involved in reaching common ground on contentious scientific issues. But there are also countless hours and late nights of ploughing through the latest research, analysing gigabytes of data and responding to thousands of comments by reviewers.

Once elected by the IPCC, authors are engaged in a tightly scheduled three-year process that encompasses multiple rounds of draft production, revision and finalization. A long consensus-finding process is needed, from multistep, worldwide reviews of report drafts to the preparation of a carefully worded summary for policy-makers that requires government approval. Headline statements generated by this process have made it verbatim into the decision documents of the international climate negotiations.

Yet scientists' work for the IPCC is voluntary, unpaid and mostly unassisted. And the burden on the scientists has become heavier with each cycle, leading some to question whether they can afford to work on future assessments.

This week, a task group on the future work of the IPCC will consider such issues at a meeting in Geneva, Switzerland (16–17 September). Before the panel starts to formulate the timeline and structure of working groups in early 2015, ahead of the sixth assessment, scientists and governments need to consider how the process can be made less burdensome for those involved. The second half of 2015 will see the election of the new IPCC leadership, who will then flesh out and implement the panel's decisions.

During our work for the IPCC, we collected many views and suggestions from colleagues on ways to improve the process. As the latest cycle ended, we surveyed the authors who report on the physical-science basis of climate change. Here we summarize their responses and outline two approaches ▶

for how we think the assessment could be improved. We call for careful evolution of the current comprehensive assessment system and greater support for participants from their host institutions. This would guarantee that the best and most robust scientific information will continue to be delivered to the climate-policy process and the public.

THE GROWING BURDEN

IPCC assessments are prepared by three working groups. The first reports on the physical-science basis of climate change; the second on impacts, adaptation and vulnerability; and the third on mitigation of climate change. To gauge climate scientists' opinions about the most recent assessment process, IPCC Working Group I surveyed its authors (see 'Author survey'). Questions focused on two issues: whether the scientific community is still able to carry out the volume of work required by the current system; and whether adjusted approaches might provide the information that stakeholders will need seven to ten years from now in a more accessible way.

More than 80% of the 172 respondents (66% of those polled) rated their overall experience as an author as very good to excellent, indicating that after 25 years physical scientists continue to strongly support IPCC assessment work. Difficulties in digesting the mountains of literature were flagged by more than 80%. More than 60% encountered hurdles when processing big data associated with the analysis of model simulations for climate projections and had trouble gaining timely access to model results[1].

We sought further opinions from participants in special sessions held at the American Geophysical Union annual meeting in December 2013 and the European Geophysical Union general assembly in April 2014. Two issues dominated: the work burden and difficulties in the transfer of assessment findings to the other IPCC working groups.

The volume of information challenges even the most enthusiastic and efficient scientists. For the fifth assessment report, Working Group I assessed more than 9,200 peer-reviewed articles and analysed more than 2 million gigabytes of numerical data[2]. Authors did this on top of their regular jobs, mostly at universities or in research laboratories. Many relied on informal help from colleagues. A further 600 contributing authors and 1,000 expert reviewers made substantial contributions.

Governments and universities want their best scientists elected to the IPCC. But those scientists need support throughout the assessment process, not just at the election stage. Institutions should reduce the administrative and teaching load of authors to free up time for their IPCC work.

IPCC authors should not, in our view, receive direct financial compensation — that

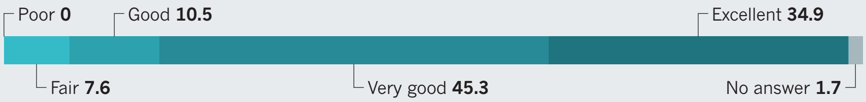
AUTHOR SURVEY

In April 2014, the co-chairs of Working Group I (WGI) of the Intergovernmental Panel on Climate Change (IPCC) invited 259 WGI coordinating lead authors, lead authors and review editors to take an online questionnaire on their experiences. Of these, 172 responded.

1 Attitude towards, and willingness to serve, the IPCC

More than 90% rated their overall experience as good or better. Meanwhile, 68% would serve again; 20% would not. The role of review editor was widely criticized as having responsibility without power.

Please rate your overall experience.

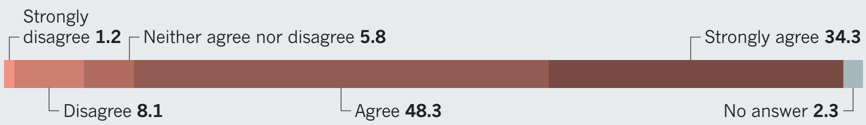


Assuming the IPCC process is unchanged would you be willing to serve again?

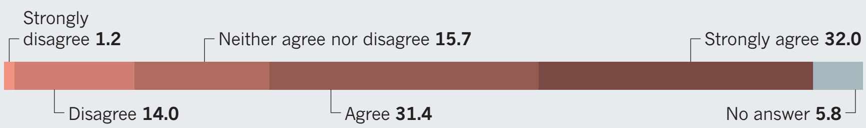
2 Workload

Since governments commissioned the first assessment report, published in 1990, the burden on the scientists has increased at an accelerated pace. A search for 'climate change' in the Thomson Reuters Web of Science yields 7,106 articles from 1900 to 2000, the time of the third assessment report. More than 110,000 articles published since 2001 include the term.

The amount of literature to be assessed was a challenge.



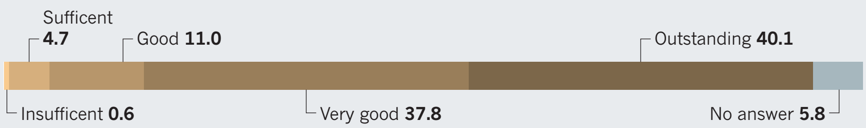
The amount of data to be processed was a challenge.



3 Assistance

The responses underline the importance of technical support units for IPCC working groups. About 80% of respondents felt that extra assistance was also necessary for those who coordinate a chapter team.

Please rate the overall support that you received from the WGI Technical Support Unit.

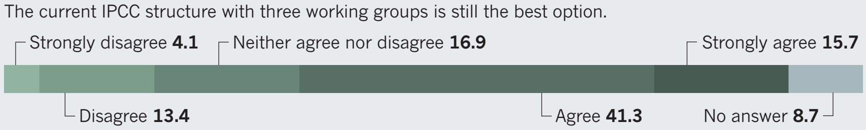


Dedicated assistance for chapter coordinators should be a standard approach in future assessments.

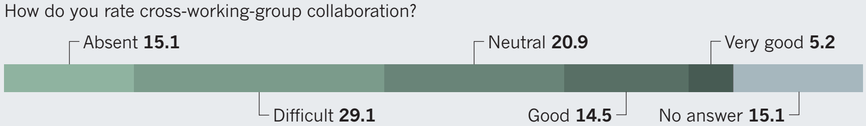
4 Working Group structure

The majority see no need to change the structure of the three IPCC working groups. They do, however, identify a deficit in collaboration between groups.

The current IPCC structure with three working groups is still the best option.



How do you rate cross-working-group collaboration?



Not all values add up to 100% because of rounding.

would risk creating conflicts of interest. But those with significant responsibilities, such as lead authors who coordinate chapter teams, work on cross-cutting issues or serve in more than one working group, should be provided with the means to hire a science assistant or postdoc for the duration of their IPCC service. The benefits justify the extra cost: the author's scientific productivity could be maintained and younger scientists can learn on the job.

The interaction between IPCC working groups has long been challenging. Different communities have differing philosophies, approaches and terminologies, and mismatched time constraints regarding, for example, the availability of model simulations for impact assessment and regional analysis. However, what has been encouraging is the experience with the recent joint-working group Special Report[3], and cross-working-group expert meetings on, for example, greenhouse-gas metrics or attributing climate change and impacts to the increase in greenhouse-gas concentrations. The structured process of a joint report requires the authors to find common ground across disciplines.

TWO OPTIONS

Here we propose for discussion two approaches for future IPCC assessments that have emerged from exchanges with colleagues, at professional meetings and from our personal experience. A requirement for any approach is that the IPCC assessment must remain rigorous, robust, comprehensive within its scope, and transparent[4,5]. Any compromise on these qualities will reduce the usefulness and jeopardize the impact of future assessments.

Extend the cycle and reduce parallel efforts. The period for an IPCC assessment could be lengthened to eight to ten years, from six. The freed-up time could be invested in collaborative work on issues that cut across working groups, such as the water and biogeochemical cycles, greenhouse-gas metrics, risk of abrupt climate change and irreversibility, ocean acidification, or regional climate change and impacts. Currently, such overlap issues are dealt with separately, resulting in parallel efforts that risk inconsistencies and the doubling up of work.

Jointly scoped 'topical assessment papers' could be written by teams collaborating across working groups. Each paper would undergo a separate expert-nomination process and a formalized expert and government review. Their length would correspond roughly to what now constitutes a chapter, about 80 pages, and they would form the building blocks of the comprehensive assessment. Production time could be flexible. Each working group would weave the topical assessment papers into its comprehensive report as it went along.

A longer cycle would also allow the working groups dealing with impacts and mitigation to start later than the others. This way, much more of the most recent results from climate-model projections would be ready for impact assessment than was the case in the fifth assessment report published in 2013–14. Towards the end of the cycle, these reports would be synchronized so that the three working groups could prepare a final, succinct synthesis report.

IPCC reports would become leaner and the topical assessment papers could respond to emerging issues. But the production process would be more complex to coordinate and thus would require careful and more extensive scoping at the start.

Cut across working-group boundaries. Collaboration between the disciplines could be intensified by a series of 'special reports' that cut across IPCC working groups. Around five such reports could be conceived for the next cycle, on topics such as observed climate change and impacts, on projections and their impacts, on scenarios and climate targets, and on the costs of climate-change adaptation and mitigation.

"The interaction between IPCC working groups has long been challenging."

Each special report would be developed under the joint responsibility of two working groups, with one leading, and include a regular scoping and expert-nomination process. Timings would be set by the availability of scientific material, for example, analysis of relevant satellite observations or climate-model simulations. A summary for policy-makers would be approved for each special report with an overarching, joint technical summary and policy summary.

The downsides of this approach include the risk of not being comprehensive and the increased management burden for the IPCC and governments.

CAREFUL EVOLUTION

Many other opinions and suggestions have been aired. Regionalization of IPCC assessments is sometimes called for to give policy-makers and practitioners more and better regional information. In our view, this approach would undermine the global character of the climate-change problem exemplified by the IPCC.

Wiki-type assessments have a modern and transparent appeal, but they lack the robustness of the formal IPCC process. Comprehensive assessments done only every ten years, but alongside annual updates on the ongoing anthropogenic climate change, would duplicate well-established efforts by others, including the American Meteorological Society[6].

All such proposals would require fundamental changes to the established and successful IPCC assessment process that has been in place since 1988. And many of these changes would, in our view, reduce the scientific rigour and comprehensiveness and thus threaten the essence of an IPCC assessment.

Ongoing negotiations on a new international climate-change agreement within the United Nations Framework Convention on Climate Change, its implementation and future adjustments, call for a continuation of comprehensive climate-change assessments by the IPCC.

Our preference is the first of the approaches presented here. Topical assessment papers would increase the responsiveness of the IPCC to emerging issues that are relevant for policy-makers, while keeping the comprehensive nature of the full assessment.

To respond to growing regional needs, the Working Group I Atlas[7] could also be extended to include quantities that would be relevant for humans and ecosystems, for example, maps of exposure and vulnerability. Ultimately, this could result in global maps of projected risks for a plethora of global-scale climate processes. Such information might be used by emerging national climate services, offering regional analyses for decision-makers, which could supplement the assessed information with their own products and databases.

Irrespective of the IPCC products — classical or new — enhanced support for scientists in responsible positions is essential for the next cycle. For the sixth assessment, the IPCC needs to consult widely and design an approach that is useful for policy-makers and feasible for scientists. ■

Thomas F. Stocker *is professor of climate and environmental physics at the University of Bern, Switzerland, and co-chair of Working Group I of the Intergovernmental Panel on Climate Change.* **Gian-Kasper Plattner** *is head of the Technical Support Unit of IPCC Working Group I at the University of Bern, Switzerland.*
e-mail: stocker@climate.unibe.ch

1. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. *Bull. Am. Met. Soc.* **93,** 485–498 (2012).
2. IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2013).
3. IPCC. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* (Cambridge Univ. Press, 2012).
4. Solomon, S. & Manning, M. *Science* **319,** 1457 (2008).
5. Stocker, T. F. *Nature Geosci.* **6,** 7–8 (2013).
6. State of the Climate 2014 in *Bull. Am. Met. Soc.* **95,** (eds J. Blunden & D. S. Arndt) S1–S257 (2014).
7. IPCC *Annex I: Atlas of Global and Regional Climate Projections* (Cambridge Univ. Press, 2013).

COMMENT


WANG DINGCHANG/XINHUA PRESS/CORBIS

Some of China's most powerful wind turbines, at the Donghaitang wind farm in Wenling.

Manufacture renewables to build energy security

Countries should follow China's lead and boost markets for water, wind and solar power technologies to drive down costs, say **John A. Mathews** and **Hao Tan**.

China's rise to become the world's largest power producer and source of carbon emissions through burning coal is well recognized. But the nation's renewable-energy systems are expanding even faster than its fossil-fuel and nuclear power. China leads the world in the production and use of wind turbines, solar-photovoltaic cells and smart-grid technologies, generating almost as much water, wind and solar energy as all of France and Germany's power plants combined[1]. Production of solar cells in China has expanded 100-fold since 2005.

As the scale of Chinese manufacturing has grown, the costs of renewable-energy devices have plummeted[2]. Innovation has played a part[3]. But the main driver of cost reduction has been market expansion. Germany and South Korea are following similar paths. In short: industrialization can go hand in hand with decarbonization.

Too many countries have yet to take notice. The United States and European Union are pursuing counterproductive policies, such as increasing trade tariffs on imported Chinese photovoltaic panels. Restricting global trade in renewable devices will only slow the rate at which costs decrease and will decelerate the world's retreat from fossil fuels.

As a result, uptake of renewable energies globally has been too sluggish to seriously reduce greenhouse gases and tackle climate change. For 15 years, countries have failed to deliver their carbon-reduction commitments under the Kyoto Protocol, hindered by the vested interests of the fossil-fuel industry and fears that the alternatives are costly.

The narrative around renewable energies needs to change. As in China, renewables must be seen as a source of energy security, not just of reduced carbon emissions. Today's discussions about energy security focus almost exclusively on maintaining access to fossil fuels. But unlike oil, coal and gas, the supplies of which are limited and subject to geopolitical tensions, renewable-energy devices can be built anywhere and implemented wherever there is sufficient water, wind and sun.

GREEN GROWTH

As the scale of manufacture and use of renewables rises, market forces will make them more accessible, affordable and

efficient. Energy policies should therefore focus on promoting manufacturing, trade and competition in low-carbon technologies, rather than supporting ever more expensive, dangerous and inaccessible fossil fuels. Emissions reductions will follow.

China generates more than 5 trillion kilowatt-hours (kWh) of electricity, about 1 trillion kWh more than the United States. China's rapid economic expansion since it joined the World Trade Organization (WTO) in 2001 has been based on fossil fuels: it consumes around 23% of the world's coal production for electricity. But fossil fuels alone cannot power the industrial growth the country needs to keep up with the West.

Since the mid-2000s, China has also pursued a low-carbon energy strategy. Investment in hydroelectric, wind, solar and nuclear-power generating facilities increased by 40% between 2008 and 2012 — from 138 billion renminbi (US$22 billion) to about 200 billion renminbi. The share of investment in fossil-fuel power facilities in China, meanwhile, fell from around 50% to 25% over the same period.

As a result, China's wind-power capacity has increased fivefold in the past four years (see 'Wind speed'). And in 2013, the generating capacity from new water, wind and solar sources exceeded[4] that of new fossil-fuel and nuclear facilities for the first time (see 'Renewables powerhouse'). Zero-carbon sources now contribute 9.6% of the energy used in China, up from 5.6% in 2000. This is a considerable achievement.

In 2013, China also hit its target — two years early — to generate almost 30% of electricity from renewables. The Chinese government aims for renewables capacity to reach 550 gigawatts (GW) by 2017, or 48% above the 2013 level. No other country is investing so much money or generating so much renewable energy.

ECONOMIES OF SCALE

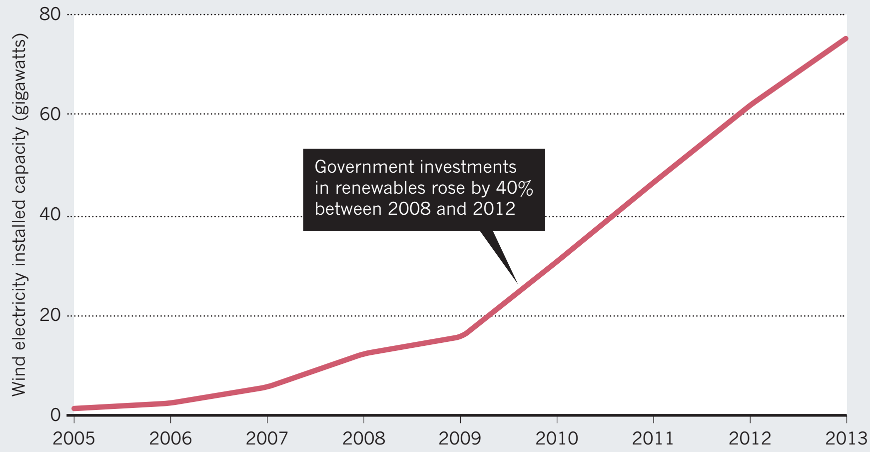
China is upgrading its power grid to accommodate power fluctuations and distributed generation for intermittent sources. In one demonstration project, the State Grid Corporation of China (SGCC) is investing 9.4 billion renminbi to integrate wind and solar-photovoltaic generation and storage devices into the main grid. The SGCC is helping to set international product standards for smart-grid elements that will underpin the export of these technologies to countries such as Brazil.

How has China's energy security improved? China became a net importer of oil in 1993, of natural gas in 2007, and of coal in 2011. Hitting its 2017 wind, water and solar power targets, we calculate, would translate into a saving of 45% on current imports of oil, coal and natural gas.

There are two keys to China's success in renewables. Focused policies drive investment in selected sectors and encourage domestic take-up by measures such as feed-in tariffs. And industrial dynamics, including economies of scale and efficiencies gained through learning, drive down unit costs as the global market expands.

WIND SPEED

Wind-power capacity has risen fivefold in China in the past four years. Turbine costs fell as the scale of manufacturing rose.



SOURCE: EIA/CHINA ELECTRICITY COUNCIL

Renewable-energy generation requires the manufacture of many components, such as wind turbines, solar-photovoltaic cells, mirrors, lenses, batteries and energy-storage systems. From 2010 to 2013, while total global photovoltaic installation more than tripled from 40 GW to 140 GW, China's installation expanded 22-fold, from 0.8 GW to 18 GW. Supplying the international market, as well as the domestic one, has helped to drive down costs of photovoltaic panels by 80% since 2008. Solar-power users around the world have benefited from lower prices.

"No other country is investing so much or generating so much renewable energy."

A few other countries are following a similar strategy. South Korea, for example, is committed to 'green growth' — expanding its smart grid and focusing its production on emerging clean sectors such as zero-emission vehicles. And Germany has been expanding its manufacture and use of solar and wind power (under its *Energiewende* energy-transition programme) since the early 2000s, with the aim of replacing its nuclear power with renewables.

The same principle of industrial-scale production established US supremacy in the automotive industry a century ago. Between 1909 and 1916, Henry Ford reduced the cost of his Ford Model T by 62%, from $950 to $360. Each year, sales doubled — from fewer than 6,000 in 1908 to more than 800,000 in 1917.

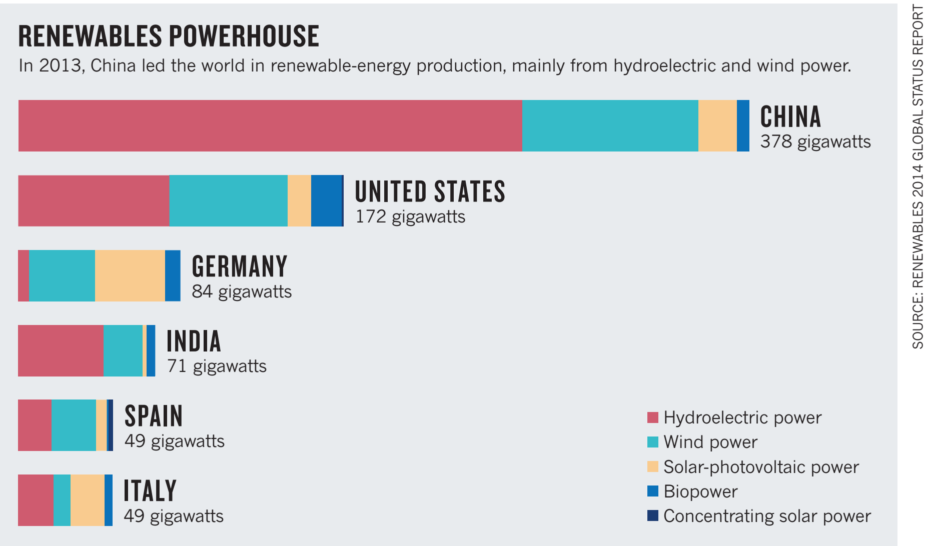
Yet US energy policy emphasizes exploiting domestic coal seam gas and shale oil, through innovations such as hydraulic fracture (fracking) and horizontal drilling. The problems of diminishing returns and environmental costs of fossil fuels remain[5]. The United Kingdom, too, is inclined to build up its supplies of coal seam gas by fracking, and to expand its fleet of nuclear reactors, a portfolio approach that will leave the country importing others' technology.

CHANGING THE CONVERSATION

Reframing the emissions debate in terms of energy security has profound implications for international negotiations under the terms of the United Nations Framework Convention on Climate Change. In December, national representatives will gather in Lima for the preparatory meeting to the Paris conference in 2015. Their agenda remains negotiating voluntary national carbon-emissions reductions, rather than promoting renewable-energy industries, as the fastest route to decarbonization.

But governments that build strong renewables sectors can achieve those emissions reductions while enhancing their energy security and building their manufacturing industries. Another advantage of the market-oriented approach is that renewables are not burdened with the task of resolving the entire climate-change problem. Few countries will be able to rely on water, wind and solar power alone, and some fossil fuels will continue to be used.

Our critics will counter that technology-based solutions raise concerns over the availability of industrial materials and land for building solar and wind devices and farms. But our calculations suggest[6] that a global renewables push for an extra 10 terawatts of power-generation capacity could be achieved on current industrial scales over the next 20 years, by which time the world energy



system would be well on the way to total conversion. Producing the extra 10 terawatts from renewables needed to transform global electric power would require more than 5 million square kilometres (about twice the size of Kazakhstan) filled with around 3 million wind turbines, 14,000 concentrated solar-power installations and 12,500 solar-photovoltaic farms. These technologies could perhaps be accommodated in the world's desert and semi-desert regions. The targets are large — but they are manageable compared with current world production levels of 1.75 billion mobile phones per year or 84 million vehicles per year[6].

TRADE SOLUTIONS

The main obstacles to expanding renewables uptake are failed policies and continuing subsidization of fossil fuels.

All governments should enlarge the market for renewable power by encouraging manufacture and trade of devices. Countries should foster export and import of renewable electric power (from, say, North Africa to Europe under the DESERTEC project, or from Mongolia to China, Japan and South Korea under the east Asian super-grid proposal). Above all, the narrow agenda that the Kyoto process has enforced needs to be broadened.

> *"These technologies could perhaps be accommodated in the world's desert and semi-desert regions."*

How? One way involves expanding free trade in renewable devices. Here, the WTO could complement the Kyoto process[7]. A preliminary agreement to free up trade in renewables was adopted by Asia-Pacific Economic Cooperation countries in 2012, and could be proposed to the WTO. A precedent exists with trade in personal computers and other information-technology products. It was expanded from a voluntary agreement to reduce tariffs, signed up to by most major industrial countries, and adopted by the WTO in 1997.

Private finance must also play a part. The Kyoto-process negotiators have so far considered that financing for climate-related initiatives should come from tax-based public finance rather than from private or even government-backed development banks. This emphasis needs to change. Green bonds lower the costs of capital and facilitate the scaling up of investments. One example is the $500-million bond issued by the Export-Import Bank of Korea last year allocated exclusively to finance green projects around the world.

China is leading the way. By placing the emphasis on production scale and market growth, it is contributing more than any other country to a climate-change solution. Its build-up of renewable-energy systems at serious scale is driving cost reductions that will make water, wind and solar power accessible to all. ■

John A. Mathews *is professor of strategic management at Macquarie Graduate School of Management, Macquarie University, Sydney, Australia.* **Hao Tan** *is senior lecturer at Newcastle Business School, University of Newcastle, Callaghan, Australia.*
e-mail: john.mathews@mgsm.edu.au

1. BP. *BP Statistical Review of World Energy 2014* 63rd edn. (BP, 2014).
2. World Energy Council *World Energy Perspective: Cost of Energy Technologies* (World Energy Council, 2013).
3. Trancik, J. E. *Nature* **507,** 300–302 (2014).
4. Mathews, J. A. & Tan, H. *Asia Pac. J.* **12,** (2014); available at http://go.nature.com/z6job5.
5. Mathews, J. A. & Reinert, E. *Futures* **61,** 13–22 (2014).
6. Mathews, J. A. & Tan, H. *J. Sustain. Energy Eng.* http://dx.doi.org/10.7569/JSEE.2014.629505 (2014).
7. Helm, D. *Nature* **491,** 663–665 (2012).

THE BRITISH LIBRARY BOARD

A landscape from Jules Verne's *Journey to the Centre of the Earth*, published 150 years ago.

SCIENCE FICTION

Verne and beyond

Danièle Chatelain and **George Slusser** explore how French science fiction grapples with Cartesian duality.

Hundreds of years before Jules Verne's heyday, France was in the vanguard of science fiction — driven by the world view of an extraordinary scientist-philosopher. René Descartes' 1644 *Principles of Philosophy* launched the paradigm of a 'clockwork' cosmos of matter and motion mastered by the rational yet metaphysical mind. Although later criticized as the "ghost in the machine", this mind–matter duality inspired a tradition of reasoned speculation about the nature of the world.

Pierre Gassendi was its fountainhead. An empiricist, he published the first data on the transit of Mercury in 1631, posited the idea of infinite space and urged open-ended investigation of the material world. Gassendi decried metaphysics, but was fascinated by Descartes' idea of a mind probing the Universe, which Enlightenment technology was then gradually revealing.

Gassendi's literary inheritor was his pupil, the dramatist Cyrano de Bergerac, who used the device of an imaginary voyage to advance the idea of empirical observation of new worlds. Bergerac's 1657 *Comical History of the States and Empires of the Moon*, which features lunar voyages propelled by rockets and dew, is often seen as the first fictional exploration of gathering and experimenting with data. Simon Tyssot de Patot expanded the field with *Voyages and Adventures of Jacques Massé* (1714). One of the first novels with a 'lost race' theme, it features living fossils such as gigantic birds surviving from prehistory — itself then a heretical concept.

French science fiction began to play more seriously with time and space with Louis-Sébastien Mercier's 1770 *Memoirs of the Year Two Thousand Five Hundred*, which treated the future as a new-found country — in this case, a Paris with functional hospitals and no beggars. Writers such as Restif de la Bretonne in *Les Posthumes* (1802) and Émile Souvestre in the 1846 *The World As It Will Be* continued to explore time travel. England had long been producing similar speculations, from Francis Godwin's *The Man in the Moone* (1638) to Daniel Defoe's *The Consolidator* (1705). But these did not struggle with the challenge recognized by the French tradition — the ambiguous role of the mind in scientific exploration.

With literary giant Honoré de Balzac, the early-nineteenth-century interest in biology and physics began to feed a substantially scientific fiction. In *The Centenarian* (1822), Balzac grapples with the quest to extend human life, much as Mary Shelley had done in *Frankenstein* four years before. But for Balzac, the quest is free of Shelley's religious and moral considerations. *The Centenarian* embraces humanity's material condition: the mind dies with the body. The protagonist keeps his body alive by using elaborate laboratory apparatus (Frankenstein has no equipment) to distil the vital fluid from other humans. And whereas Frankenstein remains an alchemist, Balzac develops a law of 'human thermodynamics' influenced by physicists Nicola Léonard Sadi Carnot and André-Marie Ampère. This dictates that every mental act of wishing or willing results in an equal, opposite and irreversible reduction of bodily resources; the only way to break this infernal circle is to import energy.

English fiction, by contrast, did not fully engage with new scientific theory and method until H. G. Wells's 1895 *The Time Machine*, which places human activity in the uncompromising perspective of evolutionary theory and features a scientist time traveller. Between Shelley and Wells, the British field was dominated by oddities such as essayist Thomas De Quincey's partly fantastical musings on his country's rush towards technological supremacy, *The English Mail-Coach* (1849).

Verne spent the latter half of the nineteenth century extending the bridge between ▶

literature and science. *Journey to the Centre of the Earth* (1864) is about scientific method and its misuses. Scientists Professor Lidenbrock and Axel enter Earth through an Icelandic crater and, after improbable adventures involving mastodons and underground oceans, are ejected through the Italian volcano Stromboli. Lidenbrock ignores data that disturb his schema. Axel is a romantic who fails to examine observable facts. Yet the book probes scientific wonder: when Axel is lost and terrified in subterranean darkness, the reader experiences awe contemplating the complete absence of light.

The French-language genre advanced significantly with the uncompromising scientific approach of J.-H. Rosny Aîné — the pseudonym of the Belgian Joseph Henri Honoré Boex. In the 1910 *Death of the Earth*, Rosny's vision of global environmental crisis is prescient. An imbalance created partly by humans turns Earth to desert. Targ, the last man, succumbs with Darwinian altruism. Realizing that carbon-based life must perish so that the iron-based Ferromagnetics can inhabit the stricken planet, he invites them to take his blood. Rosny excised the anthropomorphic from science fiction.

The 1950s and 1960s saw an invasion of space-age Anglo-American sci-fi, quickly rejected by French critics. Its main portal was *Fiction*, launched in 1953 as a French edition of the US *Magazine of Fantasy and Science Fiction*. From the outset, its editors used it as the platform for a new French sci-fi school relocating space expansionism to 'inner space' and exploring 'mind travel'. In Gérard Klein's *The Overlords of War* or Kurt Steiner's *The Scratched Record* (both 1970), time travel occurs in a vast mindscape generated by huge computers.

In French neuroscientist Jean-Pierre Changeux's scientific treatise *Neuronal Man* (1983), consciousness is linked to brain biology, breaking Descartes' duality. Yet mapping the mind in the brain is a work in progress. There remains plenty of scope for Gallic sci-fi to explore consciousness: the Cartesian ghost still lurks in the French vision of mind and matter. ■

Danièle Chatelain *is professor of French at the University of Redlands, California.* **George Slusser** *is professor of comparative literature and curator emeritus of the Eaton Collection of Science Fiction and Fantasy at the University of California, Riverside. They co-published the critical editions and translations of Honoré de Balzac's* The Centenarian *and* Three Science Fiction Novellas *by J.-H. Rosny Aîné.*
e-mail: daniele_slusser@redlands.edu

REX

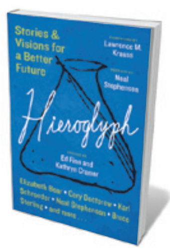
Q&A Neal Stephenson
The sci–fi optimist

Best-selling science-fiction writer Neal Stephenson's works cover everything from cryptography to Sumerian mythology. Ahead of next year's novel Seveneves, *he talks about his influences, the stagnation in material technologies, and* Hieroglyph, *the forthcoming science-fiction anthology that he kick-started to stimulate the next generation of engineers.*

What sparked your interest in science?
There were scientists in several generations of my family. My father was an electrical engineer. I grew up in the university town of Ames, Iowa, which was the best place to grow up in the history of the world, if you were a kid with an interest in science. My friends' parents had PhDs or were studying for them. Respect for science was implicit. I am drawn to 'hard' sciences because I have tools for understanding them, and it is the culture I came from.

How did you become a writer?
As a kid, I read a lot of science fiction and Classics Illustrated comics, and had a series of gifted English teachers — so it wasn't a completely alarming career choice. In college I took a mishmash of physics, geography and computer programming subjects that never added up to a marketable degree. I found myself working as a typist at the University of Iowa libraries, writing my third novel sitting on a milk crate with a fan, beer and a fancy rented typewriter. It was so hot that July that the typewriter's plastic ribbon kept sticking to its internal parts. I figured out that it only got stuck if the ribbon stood still for long enough, so I hammered the thing out. It was accepted and editor Gary Fisketjon spent a year cleaning up my "loose and baggy monster". That became my first published novel, *The Big U* (1984, Harper Perennial), a broad, science-fiction-inflected satire of college life.

How much background research do you do?
I veer back and forth between trying to do the right thing and blind panic. After *The Big U*, I thought I would write about physics. The idea was that the huge explosion in Tunguska, Russia, in 1908, was caused by a primordial singularity — a tiny black hole — popping in and out of Earth. I had a conceit that people following it put the equivalent of a bungee cord around it and got pulled out into space. I spent years writing this thing — and it was terrible. I was so scared that I had blown my chances of being a writer that I wrote another book in 30 days. That turned out to be my second published novel, *Zodiac* (1988, Atlantic Monthly).



Hieroglyph
EDITED BY ED FINN AND KATHRYN CRAMER
HarperCollins: 2014.

How does attending scientific meetings inform your writing?
I go on the spur of the moment. It is good to be in touch, to see what people are working on.

I can also get a sense of personalities and ideas — although I try to avoid focusing on specific living people in my books.

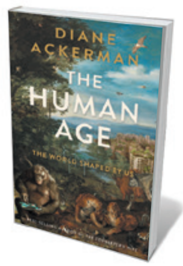
What is *Hieroglyph*?
It was born from a friendly argument with Michael Crow, president of Arizona State University in Tempe. I was complaining that progress in material technology has petered out. We have taken the creativity that went into designing rockets and channelled it into information technology (IT). A lot of bright people are dedicating their lives to inconsequential things: writing apps and so on. There is a lack of grandeur. Crow said, "It's your fault. You sci-fi writers need to give us something to work on." So the university, with my input, founded the Center for Science and the Imagination and launched Project Hieroglyph as an online forum where science-fiction authors could write in an optimistic vein, positing attainable technologies for young engineers. The collection *Hieroglyph*, out this month, showcases work by 20 visionaries, including astrophysicist and award-winning writer Gregory Benford, and science-fiction authors Cory Doctorow, Elizabeth Bear and Bruce Sterling. My contribution is 'Atmosphaera Incognita', about the construction of a 20-kilometre steel tower and the resulting adventures.

What do you think about the trend for apocalyptic science fiction?
In the 1950s we could see that we have a rocket and if we build a bigger rocket, we could go to the Moon. But with advances in nanotechnology and IT, there are many imponderable outcomes. It is easier to predict a gloomy one. But that has led to lazy, derivative, predictable stories, especially on television and in movies.

What do you think about the rise of anti-science feeling in the United States?
It is a surprise to me. Growing up in Ames, I went to a Methodist church filled with professors who never would have questioned the validity of evolution. I think a lot of opposition to global warming and evolution is not about science. The majority of people who identify themselves as global-warming sceptics, for example, do believe it is happening. But they think that admitting that will open the door to excessive regulation by the government. They don't come from the scientific community, where it is important to say what you mean. They come from a political community, where what really matters is the final outcome. I think it's self-destructive in the long run — people who refuse to face reality are infantilizing themselves. ■

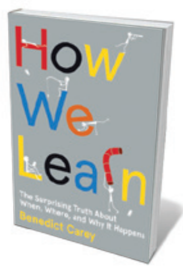
INTERVIEW BY ZEEYA MERALI

Books in brief

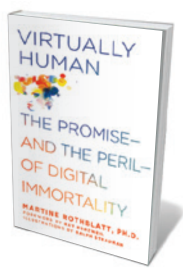


The Human Age: The World Shaped By Us
Diane Ackerman W. W. NORTON *(2014)*
The incisive yet optimistic science writer Diane Ackerman slices into the chaotic age of turbocharged technology and environmental crisis that we call the Anthropocene. She zips from deep history to speculative futures to contextualize snapshots of our vivid, frenetic present. We meet an ocean-column farmer and an orang-utan wielding an iPad; consider cross-border wildlife corridors and invasive species; wonder at the human microbiome and printed drugs. As Ackerman deciphers our grave new world, one message reverberates — that we "still and forever remain a part of nature".

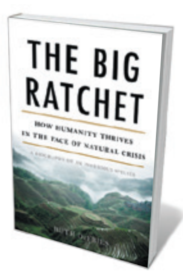
A Buzz in the Meadow
Dave Goulson JONATHAN CAPE *(2014)*
In 2003, leading bee researcher Dave Goulson bought a run-down farm in France. His aim was to provide a haven for the insects he has devoted his life to studying, notably the bumblebee. He writes beautifully of the panoply of creatures — from deathwatch beetles to dragonflies — that often pass unnoticed under our noses. But for all its easy charm, Goulson's account is permeated with awareness that biodiversity is now often confined to managed sanctuaries. What begins as a scientific rural idyll becomes a journey into the imperilled territory of Rachel Carson's *Silent Spring* (Houghton Mifflin, 1962).



How We Learn: The Surprising Truth About When, Where, and Why It Happens
Benedict Carey RANDOM HOUSE *(2014)*
Learn how to learn, enjoins science journalist Benedict Carey in this tour of past and present research on the process. Hard graft is just part of the package; what is key, Carey argues, is exploiting the brain's quirks. He lays bare the biology, cognitive science and "ways to co-opt the subconscious mind" that ensure mental labour becomes ingrained. Carey is an adroit guide to techniques for comprehension and retention, whether exploring the value of forgetting, distraction and interruption, or examining the power of studying in varied venues.



Virtually Human: The Promise — and the Peril — of Digital Immortality
Martine Rothblatt ST MARTIN'S PRESS *(2014)*
In this explication of cutting-edge artificial intelligence, technologist Martine Rothblatt argues that software brains will "express the complexities of the human psyche, sentience, and soul" surprisingly soon. Aeroplanes, she notes, lack the complexity of birds but still fly; similarly, cyber-doppelgängers or "mindclones" will emerge when symbol-association software is combined with personal information gathered on social media ("mindfiles"). Rothblatt lays out a serious analysis of the ethical and scientific implications.



The Big Ratchet: How Humanity Thrives in the Face of Natural Crisis
Ruth DeFries BASIC BOOKS *(2014)*
Vastly boosted agricultural production and cheaper food have driven today's human boom — the "big ratchet", or explosion in population over the past six decades — argues environmental geographer Ruth DeFries. Now, we are embarking on the vast experiment of feeding today's 7-billion-plus people, with no sure outcome. DeFries unpicks the historical patterns to parse the uneasy equation of people and food — our most powerful link with nature. **Barbara Kiser**

Correspondence

Frame retractions so they hold firm

The retraction last month (see *Nature* **512,** 338; 2014) of the paper 'Generation of pluripotent stem cells from adult human testis' by S. Conrad *et al.* (*Nature* **456,** 344–349; 2008) has caused some confusion in the scientific community because of its ambiguous wording, which does not serve the purpose of formally amending the scientific record.

Pluripotency is a well-defined property of stem cells both *in vivo* and *in vitro* (see, for example, J. Nichols and A. Smith *Cold Spring Harb. Perspect. Biol.* **4,** a008128; 2012). However, the retraction statement refers to the cells derived in the original paper as being "pluripotent to some level", which wrongly implies that there are different degrees of pluripotency. Such scientific sloppiness is misleading and runs counter to rigorous, evidence-based presentation of results (see, for example, *Nature* **510,** 187–188; 2014 and E. Cattaneo and G. Corbellini *Nature* **510,** 333–335; 2014) .

Furthermore, it is unclear what the statement "the original conclusions are not as robust as presented in the original paper" actually means — for example, it could imply that some or all of the earlier conclusions are not entirely invalidated. In which case, we think that those details should have been specified.

Joachim Kirsch *University of Heidelberg, Germany.*
Hans Schöler *Max Planck Institute for Molecular Biomedicine, Münster, Germany.*
joachim.kirsch@urz.uni-heidelberg.de

Pregnancy: study the mother's DNA as well

Research into the effects of epigenetic changes during pregnancy on the mother's long-term health is almost non-existent. This contrasts sharply with the wealth of attention paid to such cell-heritable changes, which alter gene activity but not DNA sequence, in the fetus and placenta as a developmental origin of health and disease (see S. S. Richardson *Nature* **512,** 131–132; 2014).

The pregnant body undergoes huge changes: extensive tissue remodelling, expansion in blood volume by as much as 100%, immunological and metabolic alterations, and extensive shifts in hormone signalling. And complications such as gestational diabetes and pre-eclampsia, which both subside after giving birth, are known to increase the mother's risk of later developing type 2 diabetes (L. Bellamy *et al. Lancet* **373,** 1773–1779; 2009) or hypertension and stroke (L. Bellamy *et al. Br. Med. J.* **335,** 974; 2007). These all have big implications for public health.

We need to proceed cautiously when building causal narratives for health outcomes, and it might be hard to study epigenetic effects in mothers when few other results are available for comparison. But grant applications, hypotheses and experimental design should not be framed by the fetus alone.

Hannah Landecker *University of California, Los Angeles, USA.*
landecker@soc.ucla.edu

Pregnancy: no safe level of alcohol

In our view, Sarah Richardson and colleagues understate the risks of alcohol consumption during pregnancy (*Nature* **512,** 131–132; 2014). Fetal alcohol spectrum disorders are among the three leading causes of intellectual disability (C. O'Leary *et al. Dev. Med. Child Neurol.* **55,** 271–277; 2013).

Alcohol can disrupt brain development throughout pregnancy, often without causing the recognizable facial changes of fetal alcohol syndrome. The child can experience life-long cognitive and behavioural effects as a result (see, for example, S. N. Mattson *et al. Neuropsychol. Rev.* **21,** 81–101; 2011).

A recent meta-analysis of 34 published cohort studies has revealed an association between moderate levels of alcohol exposure *in utero* and behavioural problems during childhood (A. L. Flak *et al. Alcohol Clin. Exp. Res.* **38,** 214–226; 2014). The study authors conclude that there is no known safe amount of alcohol that can be consumed while pregnant.

Thoughtful discussion of the risks of drinking alcohol during pregnancy is likely to enhance, rather than restrain, women's freedom in the long term.

Elizabeth R. Sowell *University of Southern California; and Children's Hospital Los Angeles, California, USA.*
esowell@chla.usc.edu
Michael E. Charness *VA Boston Healthcare System; Harvard Medical School; and Boston University School of Medicine, USA.*
Edward P. Riley *San Diego State University, California, USA.*

Count the social cost of oil sands too

Efforts to eliminate carbon pollution should not divert attention from other pressing issues that have accompanied oil-sands development (see W. J. Palen *et al. Nature* **510,** 465–467; 2014), such as indigenous rights, health inequities and social problems. In Canada, for example, housing shortages, substance abuse and food insecurity have all been attributed to Alberta's large-scale oil-sands production.

Furthermore, halting production from oil sands will not solve climate or environmental problems at a stroke. In our view, a better approach would be to ban fuels that emit large amounts of carbon dioxide, sulphur dioxide and harmful gases. This moratorium might include fuels such as coal, lignite, shale gas, and oil from tar sands or shale (see also A. Leach and B. Boskovic *Nature* **511,** 534; 2014).

In summary, it is important for energy and environmental policies to be discussed alongside those that involve public health, sustainable economic development, job creation and social justice (see also T. Measham and D. Fleming *Nature* **510,** 473; 2014).

Stephanie Montesanti* *University of Calgary, Canada.*
srmontes@ucalgary.ca
**On behalf of 8 correspondents (see go.nature.com/ichex2 for full list).*

Consider human will in psychology studies

To achieve the improvements advocated by Emily Holmes and colleagues for psychological treatments (*Nature* **511,** 287–289; 2014), researchers need to conceptually link studies of specific psychiatric disorders with fundamental processes that are shared by different disorders.

Psychologists often manipulate the environment of study participants (the independent variable) to alter a person's response or behaviour (the dependent variable). For example, they might compare the effects of threatening or neutral images on a subject's physiological arousal or memory. This approach lends itself to statistical analysis of group data, but it overlooks the important point that humans already control their environment by altering their responses. An example would be an anxious person who actively seeks safety by avoiding eye contact.

Research methodologies need to take into account the fact that such negative-feedback mechanisms exert control at all levels, including physiological, psychological and social (see T. A. Carey *Lancet* **382,** 1403–1404; 2013).

Warren Mansell *University of Manchester, UK.*
Timothy A. Carey *Centre for Remote Health, Alice Springs, Australia.*
warren.mansell@manchester.ac.uk

GENOMICS

Something to swing about

The first gibbon genome to be sequenced provides clues about how genomes can be shuffled in short evolutionary time frames, and about how gibbons adapted and diversified in the jungles of southeast Asia. SEE ARTICLE P.195

MICHAEL J. O'NEILL & RACHEL J. O'NEILL

The gibbon — a singing, swinging, south-east-Asian ape of the Hylobatidae family — occupies a poorly understood branch of the primate family tree. Despite being superficially similar to monkeys, gibbons share many traits with humans: pair bonding and monogamy; the lack of a tail; the ability to walk upright on legs; and a fondness for singing. Our understanding of gibbon evolution is now set to improve because they have just joined an exclusive primate club whose members include humans, chimpanzees and orangutans. On page 195 of this issue, Carbone *et al.*[1] report that they have sequenced the genome of Asia, a white-cheeked gibbon of the *Nomascus* genus. Their analysis unveils unique genomic features that shed light on some of the mysteries surrounding the evolutionary history of this remarkable mammalian family.

The genomes of almost all eukaryotes (plants, fungi and animals) are organized into blocks of DNA that undergo periodic reorganization. These reshuffling events, which typically occur in small increments, can lead to the emergence of species with distinct variations in karyotype — an organism's chromosome structure and number. How reshuffling occurs, and what factors lead to the fixation of new karyotypes, is an abiding genetic mystery, but mobile DNA elements are thought to have a role in some genome rearrangements. These DNA sequences, which were discovered more than 50 years ago[2], can move from one location in the genome to another, often leaving a copy of themselves behind.

By comparing Asia's genome with those of three other gibbons of the genera *Hylobates*, *Hoolock* and *Symphalangus*, Carbone and colleagues dated the divergence of gibbons from the great apes at roughly 17 million years ago (Fig. 1). However, they found that in a strikingly rapid series of speciation events that spanned 2 million years or less, an ancestral gibbon quickly gave rise to the four genera of extant gibbons. Coinciding with and perhaps reinforcing this rapid speciation is an unusually fluid karyotype[3] — the chromosomes of different gibbon species are more structurally diverse than those of any of the great apes.

Carbone and co-workers suggest that the gibbon's extreme chromosomal diversity may be attributable to a family of mobile DNA elements that is not found in other primate lineages. These LINE-1-*Alu*-VNTR-*Alu*-like (LAVA) elements are named after the three distinct mobile elements from which they derive[4]. Although each of the parental elements is common to all apes, the composite is unique to the gibbon lineage, with its origin dating to the time gibbons split from the great-ape lineage.

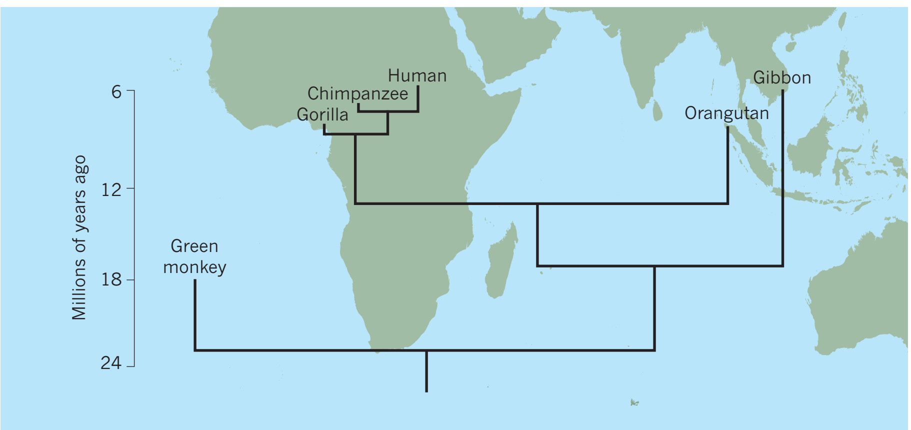


Figure 1 | Evolution of gibbons. A phylogenetic tree illustrates the evolution of great-ape species in relation to monkeys (indicated by the green monkey). Carbone *et al.*[1] estimate that gibbons separated from the great apes around 17 million years ago. The scale for divergence times is indicated on the left. The phylogeny is superimposed on a map to show where each lineage arose (although not the green monkey).

The authors' examination of the four genomes revealed the profound impact of LAVA elements on gibbon genome dynamics. In several places, insertion of LAVA elements caused premature termination of gene transcription, which might lead to the production of proteins with altered functions. The affected genes notably include several that are involved in chromosome segregation, whereby replicated chromosomes are equally distributed to progeny during cell division. Carbone *et al.* propose a scenario in which LAVA insertion results in subtly altered proteins that mildly disrupt chromosome segregation and so enhance genome plasticity. However, major alterations or a loss of function of these genes would lead to sterility or death, limiting the ability of LAVA elements to generate new chromosome arrangements.

Lending credence to the subtle-alteration model, Carbone and colleagues found that the gibbon genome contained 240 short segments (most around 150 base pairs in length) in which mutations resulting in base-pair substitutions have occurred faster than expected since separation from the great-ape lineage — a hallmark of adaptive evolution. These regions mostly lie close to the genes affected by LAVA insertions. The authors speculate that the regions may have diversified specifically in gibbons to ameliorate the detrimental effects of active LAVA elements; functional elements that modulate the impact of LAVA insertions on gene transcription were created, such as enhancers (which control gene expression from a distance)[4].

Such gene disruptions, coupled with population-size fluctuations across southeast Asia during the Miocene-to-Pliocene transition 2.5 million years ago, may have led to the extraordinary chromosomal diversity displayed in extant gibbons. Several other eukaryotic groups underwent rapid diversification in karyotype as they evolved[5], including sunflowers, Australian grasshoppers, horses and kangaroos, and the emergence of some of these species coincided with notable activity of mobile DNA. However, proof of Carbone and colleagues' subtle-alteration model will require thorough and integrative functional and evolutionary genomic analyses — a strategy that has been of great benefit to

TERRY WHITTAKER/FLPA

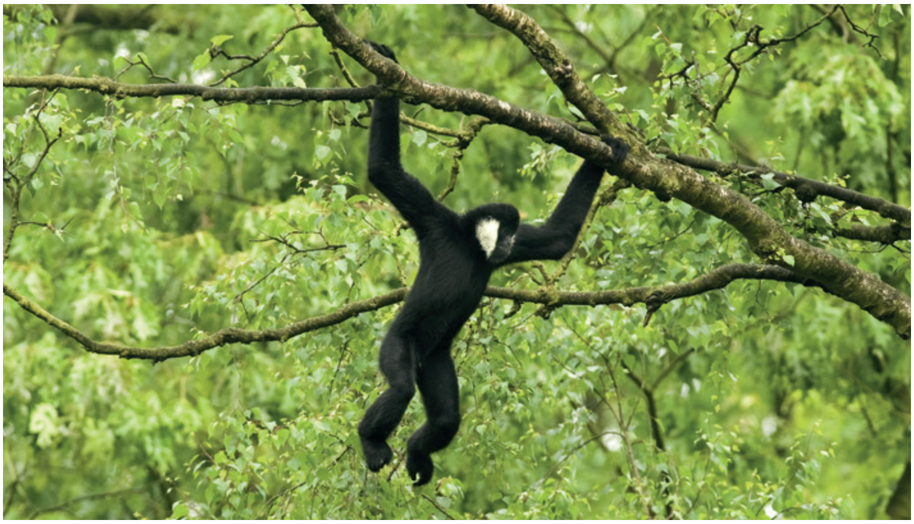


Figure 2 | King of the swingers. A gibbon brachiating through the trees.

other large-scale genomics efforts[6].

Gibbons almost fly through the trees, hitting speeds of up to 56 kilometres an hour using only their arms in a pendulum motion termed brachiation[7] (Fig. 2). The physiological features that allow this fluid and swift motion include long and powerful arms, permanently hooked hands, and a ball-and-socket wrist joint that enables swift changes in direction, even at high velocities. The authors found evidence that genes important to forelimb patterning and specialization in the gibbon have experienced a rapid evolution not seen in other primate lineages. Establishing a functional link between the adaptive evolution of these genes and gibbon brachiation is an exciting future direction — it could provide the first evidence that anatomical and locomotive specialization can be linked to specific changes at the genome level in primates.

Further exploration of the gibbon genome may shed light on other features we share with gibbons, such as their ability to sing like human operatic sopranos[8] and their penchant for walking on two legs. Publication of Asia's genome gives us something to sing about — and to swing about.

Michael J. O'Neill *and* **Rachel J. O'Neill** *are at the Institute for Systems Genomics and Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06269, USA.*
e-mails: michael.oneill@uconn.edu; rachel.oneill@uconn.edu

1. Carbone, L. *et al. Nature* **513,** 195–201 (2014).
2. McClintock, B. *Genetics* **38,** 579–599 (1953).
3. Müller, S., Hollatz, M. & Wienberg, J. *Hum. Genet.* **113,** 493–501 (2003).
4. Carbone, L. *et al. Genome Biol. Evol.* **4,** 648–658 (2012).
5. King, M. *Chromosomal Speciation Revisited (Again). Species Evolution. The Role of Chromosome Change* (Cambridge Univ. Press, 1993).
6. The ENCODE Project Consortium. *PLoS Biol.* **9,** e1001046 (2011).
7. Michilsens, F., Vereecke, E. E., D'Août, K. & Aerts, P. *J. Anat.* **215,** 335–354 (2009).
8. Koda, H. *et al. Am. J. Phys. Anthropol.* **149,** 347–355 (2012).

MICROBIOLOGY

Bacteria get vaccinated

Infection by defective bacterial viruses that cannot replicate has now been found to be the key feature enabling bacteria to rapidly develop adaptive immunity against functional viruses.

RODOLPHE BARRANGOU & TODD R. KLAENHAMMER

Surviving viral infections is a necessary ability for most life forms. Adaptive immunity, in which invasive elements are captured by the cell, allowing subsequent recognition and destruction of related viruses, is crucial for overcoming such infections. As such, adaptive immunity drives evolution, selection and fitness. Although the antibody–antigen basis of mammalian adaptive immunity has been extensively characterized, its counterpart in archaea and bacteria — CRISPR–Cas immune systems — remains largely mysterious. Writing in *Nature Communications*, Hynes *et al.*[1] describe how defective virus particles trigger immunization events by CRISPR–Cas systems, conferring adaptive immunity in the bacteria against related functional viruses.

CRISPR–Cas systems have two components: DNA sequences comprised of clustered regularly interspaced short palindromic repeats (CRISPR), and CRISPR-associated sequence (Cas) endonuclease enzymes. Typically, immunity arises when invasive genetic elements (for example, DNA injected into the cell by bacterial viruses called bacteriophages, or phages) are incorporated into the genome as 'spacers' between CRISPR sequences[2]. Subsequent transcription of the CRISPR array containing the incorporated spacers leads to the production of small interfering CRISPR RNAs[3], which guide Cas enzymes to target and cleave DNA sequences that are complementary to the spacers[4–6]. This adaptive immune system has proven to be effective against phage predation in dairy starter cultures, which are widely used in yogurt and cheese manufacturing[2].

Although it has been established that the uptake of viral DNA into the host genome drives CRISPR-based immunity[5], little is known about how bacteria sample the genomes of phages for spacer acquisition, or the dynamics of the immunization process, which can be thought of as a bacterial 'vaccination' against phages. Phages typically take over the molecular machinery of their host within minutes, and the ability of bacteria to mount a quick adaptive immune response has remained enigmatic.

Hynes *et al.* first exposed bacterial cells to defective phages. These were produced either by exposing phages to ultraviolet (UV) radiation or by using virulent phages that are susceptible to a restriction–modification (RM) system that uses restriction enzymes to cleave phage DNA after injection into the host. In both cases, the defective phages can inject DNA into the host, but cannot replicate. DNA injections by cleavage-sensitive phages result in a montage of phage DNA fragments in infected cells. Irradiation-weakened phages inject and present non-replicative DNA, which can potentially be sampled and acquired by CRISPR arrays.

The authors searched for surviving host bacteria, and found that survivors had acquired additional spacers in CRISPR sequences, an indication that the phage DNA was accessible to the CRISPR adaptation machinery (Fig. 1). Although most of the cells died, a fraction of the infected population captured phage-genome pieces in CRISPR sequences. Specifically, the presence of UV-inactivated and RM-susceptible viruses increased the generation of vaccinated bacteria by three- to fourfold and tenfold, respectively, when compared with the presence of functional phages. This

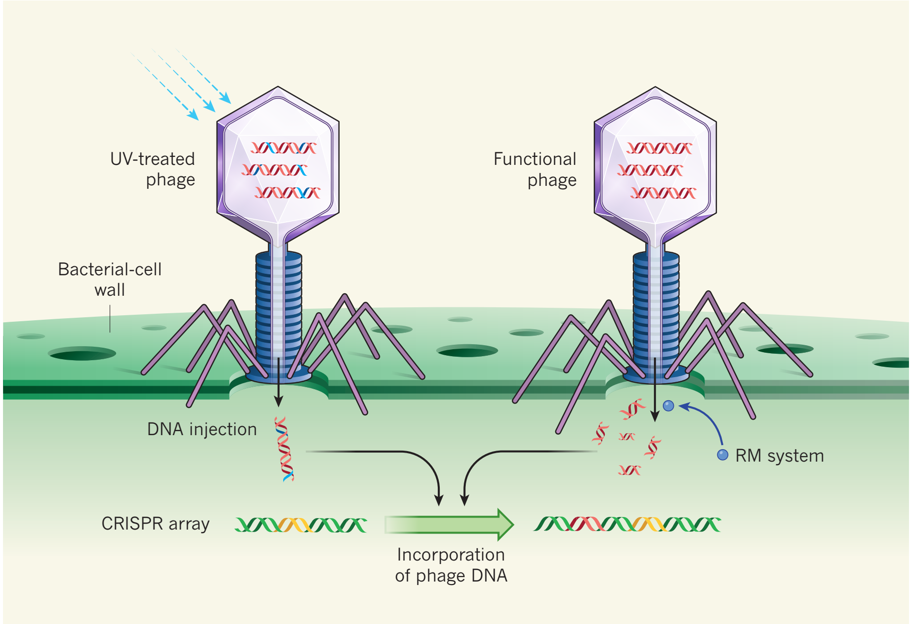


Figure 1 | Immunity through attenuated viruses. Hynes *et al.*[1] exposed bacteria to phage viruses that were unable to replicate, either because their DNA had been mutated by exposure to ultraviolet (UV) light or cleaved into fragments by the bacterium's restriction–modification (RM) enzyme system. The authors report that viral DNA fragments (red) from replication-deficient phages can be incorporated into CRISPR arrays containing spacer sequences from previous immunization events (yellow), so, through the CRISPR–Cas immune system, newly incorporated spacers provide the cell with sequence-specific adaptive immunity against related functional viruses.

implies that replication-deficient viruses drive immunization.

Next, Hynes and colleagues used a 'double viral challenge' scheme, in which bacterial hosts were concurrently infected with defective and functional phages, to demonstrate that non-replicating viruses can be used to trigger vaccination against a distinct but similar family of functional phages. This test showed that most of the vaccination events that protect the cells from the functional phages arise from the defective phage populations. The authors report a direct correlation between the proportion of replication-deficient phages used in the challenge and the number of vaccination events, compared with a challenge using only functional phages. This is reminiscent of the use of attenuated viruses and bacteria for human vaccination against pathogens.

This study provides crucial proof of concept that defective viruses can be used to trigger immunization events through CRISPR–Cas systems. Although the use of attenuated viruses for vaccination is not new, the finding that inactivated viruses can trigger CRISPR-dependent adaptive immunity in bacteria has practical implications. Thus far, analysis of CRISPR–Cas systems in general, and adaptive spacer acquisition in particular, has been hampered by the limited set of available CRISPR model systems able to acquire spacers (as opposed to just targeting and cleaving nucleic acids). We anticipate that the use of attenuated viruses will allow researchers to expand the effectiveness of CRISPR immunization, and to use CRISPR–Cas in bacteria in which the system was previously deemed to be inactive. Future studies should also establish whether sampling of chromosomal DNA and plasmid DNA (small, non-chromosomal circular DNA molecules found in bacteria and archaea) follows the same molecular rules as viral DNA.

With an increased pool of active CRISPR–Cas systems, more Cas-based molecular machines could be repurposed for biotechnological applications, such as engineering bacterial resistance to phages or plasmids, or using CRISPR–Cas technology to edit genomes and to regulate transcription in various life forms, from bacteria to animals[7,8]. Hynes and colleagues' findings will help us begin to understand the role of CRISPR in the arms race between microbial communities and their viral predators in natural habitats, and will set the stage for further applications of CRISPR–Cas systems. ■

Rodolphe Barrangou *and* **Todd R. Klaenhammer** *are in the Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State University, Raleigh, North Carolina 27695, USA.*
e-mail: rbarran@ncsu.edu

1. Hynes, A. P., Villion, M. & Moineau, S. *Nature Commun.* **5,** 4399; http://dx.doi.org/10.1038/ncomms5399 (2014).
2. Barrangou, R. *et al. Science* **315,** 1709–1712 (2007).
3. Brouns, S. J. J. *et al. Science* **321,** 960–964 (2008).
4. Marraffini, L. A. & Sontheimer, E. J. *Science* **322,** 1843–1845 (2008).
5. van der Oost, J., Westra, E. R., Jackson, R. N. & Wiedenheft, B. *Nature Rev. Microbiol.* **12,** 479–492 (2014).
6. Garneau, J. E. *et al. Nature* **468,** 67–71 (2010).
7. Hsu, P. D., Lander, E. S. & Zhang, F. *Cell* **157,** 1262–1278 (2014).
8. Sander, J. D. & Joung, J. K. *Nature Biotechnol.* **32,** 347–355 (2014).

ATMOSPHERIC CHEMISTRY

No equatorial divide for a cleansing radical

A constraint on the global distribution of the elusive hydroxyl radical takes us a step closer towards understanding the complex, interdependent factors that control the levels of this atmospheric cleanser. SEE LETTER P.219

ARLENE M. FIORE

The atmosphere is cleansed of many air pollutants and some greenhouse gases when these compounds react with the hydroxyl free radical (OH) in the troposphere, the lowermost layer of the atmosphere. Reactions with OH also prevent some substances from destroying the stratospheric ozone layer. Since the OH molecule was first confirmed to exist in the troposphere more than 40 years ago[1], a clear depiction of its spatial distribution has been elusive. Roughly speaking, the more OH that is available near pollutant sources, the faster are pollutants removed from the atmosphere, preventing their transport to other atmospheric regions. On page 219 of this issue, Patra *et al.*[2] report that there is little difference in OH abundance in the Northern and Southern hemispheres — in stark contrast to what is currently simulated by global atmospheric-chemistry models.

The OH radical reacts with other species in the atmosphere within 1 second, making its direct measurement a technical feat — one that is not possible on the time and space scales necessary to constrain its global distribution. The global mean concentration of OH can instead

be derived by analysing measured abundances of proxy compounds that are emitted to the atmosphere and removed primarily by reaction with OH, provided that the magnitude and spatial distribution of the proxy emissions are accurately known. However, long-standing discrepancies exist between the variability of global mean OH concentrations inferred from proxies and those estimated using computational models[3,4].

The best proxy for inferring global OH abundance and variability is methyl chloroform. This anthropogenic compound was once used as a solvent, and depletes ozone in the stratosphere. Emitted mostly in the Northern Hemisphere, methyl chloroform abundances are measured around the world by two monitoring networks. Aircraft flights have also sampled latitudinal variations over the Central Pacific Ocean.

Hemispheric differences in the abundances of methyl chloroform contain signatures of the hemispheric OH ratio — the ratio of annual mean OH concentration in the Northern Hemisphere to that in the Southern Hemisphere. More specifically, the differences reflect the combined influence of the amount and location of the proxy's emissions, its transport between the hemispheres and its temperature-dependent chemical depletion by OH. On the basis of observations of methyl chloroform, previously reported modelling[5] inferred hemispheric symmetry in OH concentrations. But this was not conclusive because of the uncertainties associated with methyl chloroform emissions, and because the model did not adequately represent transport across the Intertropical Convergence Zone (ITCZ) — a meteorological barrier to interhemispheric air exchange.

The production of methyl chloroform was banned by the Montreal Protocol, so only residual emissions remain. Because the rate at which it is lost from the atmosphere through reactions with OH now exceeds the emission rate, methyl chloroform can be used to infer OH levels more accurately than ever before[3]. Indeed, Patra *et al.* find that hemispheric differences in methyl chloroform abundances are no longer sensitive to uncertainties in the spatial patterns of methyl chloroform emissions.

Using differences between methyl chloroform concentrations measured at sites in the Northern and Southern hemispheres, the authors minimized the combined uncertainty associated with the global mean OH concentration and total methyl chloroform emissions. They also concluded that transport across the ITCZ in their model is accurate because it matches observed distributions of sulphur hexafluoride — an unreactive compound whose sources are better known than those of methyl chloroform and which can therefore be used to infer such transport. Having thus minimized potential sources of uncertainty, the researchers imposed spatially distinct distributions of OH on the model, and demonstrated a strong relationship between their simulated hemispheric OH ratio and the calculated difference in hemispheric abundances of methyl chloroform.

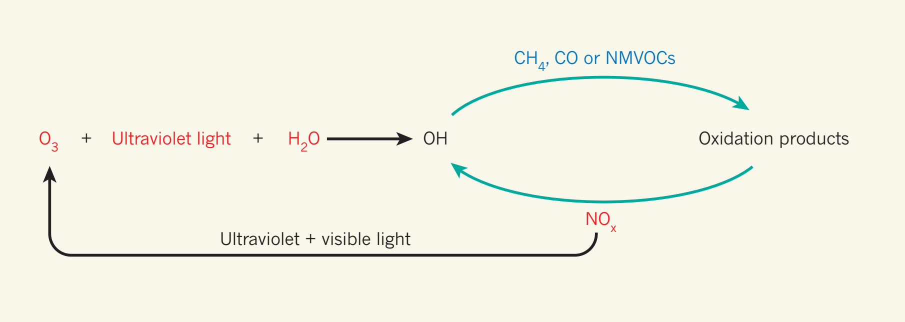


Figure 1 | Sources and sinks of hydroxyl (OH) radicals. This highly simplified scheme depicts factors that control the abundance and spatial distribution of OH in the troposphere; sources of OH are shown in red, and sinks in blue. The main pathway for OH formation is the absorption of ultraviolet light by tropospheric ozone (O_3) and its subsequent break-up in the presence of water vapour. OH radicals may then react with compounds such as methane (CH_4), carbon monoxide (CO) and non-methane volatile organic compounds (NMVOCs) to form oxidation products. Some of those products serve as extra sinks for OH before their eventual loss from the atmosphere. Others can react with nitrogen oxides (NO_x) in the presence of sunlight (at wavelengths of less than about 420 nanometres) to produce more OH and O_3. The relative importance of each source or sink varies spatially and temporally.

They found that the best match between measured and modelled interhemispheric differences in methyl chloroform levels occurs for roughly equal OH abundances in the two hemispheres. This finding directly conflicts with estimates from current global models of atmospheric chemistry, which consistently simulate higher OH levels in the Northern Hemisphere[4]. It should be noted, however, that there are open questions associated with those models: they estimate global annual mean OH concentrations that differ by ±25%, and calculate opposing responses to identical changes of anthropogenic emissions[4,6].

Although Patra and co-workers provide an invaluable service by pinning down the hemispheric OH ratio, this finding offers little insight into the complex, interdependent processes that shape OH distributions and their temporal evolution. For example, the authors say that overestimates of OH levels in the Northern Hemisphere reflect the tendency of models to calculate higher concentrations of tropospheric ozone — the main OH source — than are actually observed in that region[7]. But because ozone production in the troposphere requires OH, the calculated high ozone levels could be just another symptom of a common underlying problem: incomplete representation of some atmospheric chemical and physical processes that shape OH and ozone distributions.

Tropospheric ozone generation occurs when ultraviolet radiation splits apart ozone in the presence of water vapour. But OH can also be regenerated when methane and certain other compounds react in the presence of sufficient concentrations of nitrogen oxides — pollutants emitted by cars and smokestacks (Fig. 1; see also ref. 8). The global mean abundance and distribution of OH thus represent the net summation of these photochemical processes over myriad local environments, each of which may be dominated by a different, often poorly constrained, factor. It is changes in these individual factors that determine the evolution of OH.

Atmospheric-chemistry models are our best tools for estimating the evolution of the atmosphere's self-cleansing capacity, and for evaluating the global impacts of societal choices regarding emissions of air pollutants, greenhouse gases and ozone-depleting substances. But observational constraints with which to test these models are fairly limited. Patra and colleagues' study provides a prime example of how observational constraints can be derived, with immediate ramifications for those working in the field. For instance, the IGAC/SPARC Chemistry-Climate Model Initiative (CCMI) is coordinating a set of simulations using models that serve as workhorses for projecting the evolution of OH in the atmosphere. These simulations should be scrutinized for clues to the key processes determining the parity of hemispheric OH concentrations.

More broadly, Patra and co-workers' study implies that analysing large sets of model simulations, for example those produced through efforts such as the CCMI, can reveal clear relationships between an uncertain model parameter and a directly observable quantity. New observation-derived constraints are sorely needed to work out the complex chemical and physical processes that continuously remove harmful substances from the atmosphere. ■

Arlene M. Fiore *is in the Department of Earth and Environmental Sciences, Columbia University, and Lamont-Doherty Earth Observatory, Palisades, New York 10964-1000, USA.*
e-mail: amfiore@ldeo.columbia.edu

1. Levy, H. II *Science* **173,** 141–143 (1971).
2. Patra, P. K. *et al. Nature* **513,** 219–223 (2014).
3. Montzka, S. A. *et al. Science* **331,** 67–69 (2011).
4. Naik, V. *et al. Atmos. Chem. Phys.* **13,** 5277–5298 (2013).
5. Krol, M. C. & Lelieveld, J. *J. Geophys. Res.* **108,** 4125 (2003).
6. Voulgarakis, A. *et al. Atmos. Chem. Phys.* **13,** 2563–2587 (2013).
7. Young, P. J. *et al. Atmos. Chem. Phys.* **13,** 2063–2090 (2013).
8. Wennberg, P. O. *Nature* **442,** 145–146 (2006).

SUSTAINABLE DEVELOPMENT

The promise and perils of roads

A global map of the potential economic benefits of roads together with the environmental damage they can inflict provides a planning tool for sustainable development. SEE LETTER P.229

STEPHEN G. PERZ

Roads are seen as necessary for economic development the world over. The process by which roads are planned and built, and their impacts on affected regions, are also similar regardless of where this happens. Governments routinely plan roads without adequate consultation with local people, and construction often goes ahead with insufficient attention to minimizing the environmental effects. A mix of unexpected and unhappy outcomes then ensues, and road-building advocates are criticized for making unrealistic promises about the economic benefits and for ignoring problems such as environmental damage. There remains a need to improve the planning of roads around the world. On page 229 of this issue, Laurance *et al.*[1] take a major step towards addressing this need by presenting global maps of the potential economic and ecological consequences of future roads.

There are many scientific papers on the impacts of roads, and they draw very different conclusions. Economists have consistently documented the fact that new infrastructure fosters economic growth and reduction in poverty[2]. By contrast, ecologists have compiled a long list of environmental problems ranging from habitat degradation to species extinction[3]. Social scientists have shown that roads often cause land-use conflicts and worsen social inequality. Nonetheless, governments focus on the economic importance of roads, and populations with poor infrastructure are demanding improved access to social services and urban markets. But the reality of road impacts is decidedly mixed[4], and debate about building new infrastructure has intensified in recent years.

In this context, Laurance *et al.* provide important planning tools. The authors integrated global data sets to devise a map containing both an 'environmental-values' layer, measured in terms of the presence of protected areas, the value of various ecosystem services and biodiversity (especially of rare animal species), and a 'road-benefits' layer that estimates the potential economic benefits of new or improved roads in terms of increasing agricultural productivity and sales volume.

A key contribution of these maps is the ability to overlay them in geographic information systems to create a global planning map that identifies regions of varying potential for economic benefits and ecological damage following road building (Fig. 1). For planning purposes, three main types of area of interest emerge from these overlays: those with potentially high economic benefits, those at risk of potentially high ecological damage and those with both. The policy prescriptions are straightforward for the first two: build roads where the potential economic benefits are high and avoid them where the potential ecological damage is substantial.

But the challenge resides in the 'conflict zones' identified by Laurance and colleagues, where there is high potential for both economic benefit and ecological damage. As the authors note, these zones are key sites for the implementation of alternative policies — that is, something other than more road infrastructure is needed to solve the riddle of sustainable development in these areas. An array of policy alternatives already exists that may provide economic benefits without causing ecological damage, ranging from ecotourism to sustainable resource extraction to payments for ecosystem services.

Laurance and colleagues' map raises two main issues. First, it offers worldwide coverage, which means that it is based on a variety of data sources. Data quality is highly variable between countries, and this may have introduced biases in the findings. Their study is nonetheless helpful because it can serve as a point of departure for a broader effort to improve such maps for planning purposes. This amounts to a clarion call for the creation of an international scientific network focusing

STEPHEN G. PERZ

Figure 1 | Economics versus environment. The Interoceanic Highway in Peru ascends from the Amazon lowlands to the Andean highlands, crossing several rivers and highly biodiverse ecosystems. The highway corridor, part of the Initiative for the Integration of Regional Infrastructure in South America, is intended to expand commerce, but is facilitating illegal gold-mining, timber extraction and drug trafficking. It is an example of the 'conflict regions' identified by Laurance *et al.*[1], where road building is associated with both high potential economic benefits and great potential for environmental damage.

on roads, like the networks that already exist for land and climate science. If you think you can produce better maps of road impacts, step forward: Laurance *et al.* have placed their data products online (www.global-roadmap.org).

The second issue concerns policy initiatives to improve global road planning. Multilateral development banks fund roads to promote economic growth; in the same vein, governments build roads to support economic goals, although they also use roads for geopolitical purposes, such as securing national borders. Whether roads are built to expand commerce or improve security, a global plan for road building might be interpreted as an imposition on the priorities of sovereign countries. In particular, the conflict zones identified by Laurance *et al.* are mostly in poor countries — citing the road-planning map and telling those countries not to build roads is hardly going to be popular.

Thus, there is a need for clarity about the purpose of such maps. A global road plan is not intended to 'keep developing countries poor', but rather to highlight the costs as well as the benefits of building roads, in order to motivate a discussion of policy alternatives for sustainable development. This carries implications for the funding priorities underlying bank loans and development assistance. In cases where roads will probably cause ecological damage, governments can cite global road-planning maps to argue for policies that invest in alternative strategies for development. ■

Stephen G. Perz *is in the Department of Sociology and Criminology & Law, University of Florida, Gainesville, Florida 32611, USA.*
e-mail: sperz@ufl.edu

1. Laurance, W. F. *et al. Nature* **513,** 229–232 (2014).
2. Straub, S. Policy Research Working Paper No. 4460 (World Bank, 2008).
3. Forman, R. T. T. *et al. Road Ecology: Science and Solutions* (Island, 2002).
4. Perz, S. G. *et al. Reg. Environ. Change* **12,** 35–53 (2012).

This article was published online on 27 August 2014.

OCEANOGRAPHY

What goes down must come up

A compilation of high-resolution measurements of ocean mixing collected over the past three decades reveals how deep ocean waters return to the surface — a process that helps to regulate Earth's climate.

RAFFAELE FERRARI

Deep ocean circulation is fed by waters that become dense enough to sink into the ocean abyss in the North Atlantic Ocean and the Southern Ocean around Antarctica. These waters carry dissolved carbon away from the atmosphere and into the deep ocean, thereby playing a crucial part in modulating Earth's carbon budget and climate. Despite theoretical and observational efforts dating back to the beginning of the twentieth century, we are still struggling to understand how and where these waters return to the ocean surface — in other words, we know how ocean carbon is 'breathed in', but are still trying to figure out how it is 'exhaled'. Writing in the *Journal of Physical Oceanography*, Waterhouse *et al.*[1] report remarkable progress in resolving this long-running detective story.

The first quantitative hypothesis for the return pathway of high-latitude waters was proposed in a seminal paper[2] by the oceanographer Walter Munk in 1966. He speculated that dense bottom waters are mixed back up to the surface by breaking internal waves. To explain what this means, picture the ocean as a layer cake with colder — and therefore denser — layers at the bottom, and progressively warmer and lighter layers stacked on top. Internal waves are oscillations of these layers, analogous to the more familiar ocean waves that we see at the surface. Occasionally, internal waves overturn and break, much like surface waves on a beach, thereby mixing water from a denser layer into a lighter one and raising the potential energy of the ocean.

Direct *in situ* measurements in the years following Munk's paper, however, failed to detect enough mixing to bring back to the surface all of the high-latitude waters that sink into the abyss (calculated to require rates of approximately 30×10^6 cubic metres per second)[3].

Lacking alternative theories for the return pathway of bottom waters, oceanographers speculated that their measurements sampled areas of weak mixing and missed hotspots of intense mixing. An oceanographic gold rush to find the 'missing mixing' ensued.

Munk and fellow oceanographer Carl Wunsch quantified the amount of missing mixing on a global scale[4] in 1998. They estimated that potential energy had to be supplied at a rate of approximately 0.4 terawatts (1 terawatt is 10^{12} watts) to continuously lift dense bottom waters to the ocean surface. During an internal-wave-breaking event, about 20% of the wave energy is converted into potential energy and lifts fluid, with the rest being dissipated by inconsequential small-scale motions. Internal waves would thus have to be generated at a rate of approximately 2 TW to mix bottom waters back to the surface.

At that time, it was thought that internal waves were mainly generated by variable surface wind at a rate of less than 1 TW. Munk

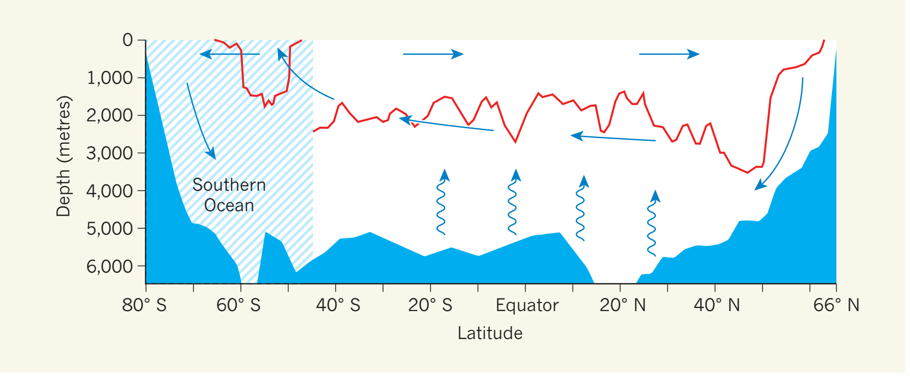


Figure 1 | An emerging model of deep-ocean circulation. Dense waters sink into the abyss at high latitudes north and south (downward arrows). The bottom waters are lifted up to depths of about 2,000 metres by mixing processes (wiggly arrows), and return to high latitudes at these intermediate depths, eventually rising to the surface via the Southern Ocean (southward and upward pointing arrows), closing the circulation loop; shading indicates the extent of the Southern Ocean. Horizontal arrows at the surface indicate the path of waters back to high latitudes. North of the Southern Ocean, the red line indicates the heights of the tallest topographic features below which mixing is strong. In the Southern Ocean, the red line indicates the topography of the Drake Passage, to illustrate topography at latitudes at which deep water is pulled to the surface by winds (the Roaring Forties). Solid blue regions indicate the deepest points at each latitude, based on a 0.25°-resolution bathymetry data set. Waterhouse *et al.*[1] confirm that this scenario is consistent with available observations of ocean mixing. (Adapted from ref. 9.)

and Wunsch[4] suggested, and later work confirmed[5], that internal waves are also generated by tidal forcing at a rate greater than 1 TW. More recently, it was shown that another roughly 0.5 TW is supplied by large-scale currents impinging on the bottom topography[6]. But just as global estimates of internal-wave generation finally seemed to be coming close to the approximately 2 TW required, *in situ* observations showed that internal waves tend to break close to ocean-bottom topography (the equivalent of beaches for surface waves), thus confining mixing to within a few hundred metres of the ocean bottom. So although the energy to support mixing was no longer lacking, the mixing was not delivered uniformly throughout the water column, as was needed to lift waters back to the surface.

The final piece of the puzzle was anticipated in 1998, when another seminal paper[7] pointed out that most of the ocean waters above depths of 2,000 m come to the surface in the Southern Ocean, where winds known as the Roaring Forties, blowing around Antarctica, pull them to the surface along surfaces of constant density. The uplift process therefore requires no mixing. Only in the past few years have oceanographers been able to integrate Munk's hypothesis with the discovery of uplift in the Southern Ocean. The emerging view is that mixing brings bottom waters in all oceans up to about 2,000 m, the characteristic depth of the most prominent oceanic topographic features. The waters then flow at approximately the same depth all the way to the Southern Ocean, where the Roaring Forties lift them to the surface (Fig. 1).

In this new scenario, the potential energy required from mixing is about half that estimated by Munk and Wunsch (the ocean is on average about 4,000 m deep, and mixing lifts the waters up to only half that depth), and it needs to be supplied in the bottom 2,000 m, the characteristic height of the major ocean ridges and sea mountains. Thus, there is no shortage of energy to support mixing, and the mixing is delivered close to the bottom topography, where it is needed. Problem solved? Not quite. *In situ* observations show that the intensity of bottom mixing is highly variable, being strong where topography is rough and bottom flows are fast, and weak elsewhere. Mapping this heterogeneity on a global scale is the next challenge in the quest to track the return journey of abyssal waters to the surface.

Enter Waterhouse *et al.*[1], who have gathered the largest compilation of *in situ* measurements of mixing so far, using them to test whether the new scenario is consistent with all available observations. They confirm that internal waves are indeed generated along the major ridges and sea mountains in the Atlantic, Pacific and Indian oceans. Most importantly, they show that about 70% of the waves break close to the ocean bottom, whereas the remaining 30% propagate away from their generation sites and end up breaking against the continental slopes. They conclude that abyssal waters make their way to the surface along the steep slopes of mid-oceanic ridges and continents, where mixing is strong.

The authors did not address the question of whether mixing is confined to depths below approximately 2,000 m — instead, they lumped together all measurements below 1,000 m. Future work must address this, because the answer is crucial for understanding and modelling the partitioning of carbon between the atmosphere and oceans. It was recently suggested[8] that the drop in atmospheric carbon dioxide concentrations recorded in ice cores from glacial periods is connected to the vertical profiles of ocean mixing. In the present climate, abyssal waters release carbon to the atmosphere when they return to the surface in the Southern Ocean. But in glacial climates, a large fraction of the Southern Ocean was covered by ice, thus trapping carbon in the ocean. This trapping was possible because strong mixing was confined to the ocean bottom, and waters could not be lifted to the surface at ice-free latitudes. Similarly, the present vertical profile of mixing will control the long-term rate (on millennial timescales) at which the ocean takes up the anthropogenic carbon we are releasing into the atmosphere. ■

Raffaele Ferrari *is in the Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Oceanography, Cambridge, Massachusetts 02139, USA.*
e-mail: rferrari@mit.edu

1. Waterhouse, A. F. *et al. J. Phys. Oceanogr.* **44,** 1854–1872 (2014).
2. Munk, W. H. *Deep Sea Res. Oceanogr. Abstr.* **13,** 707–730 (1966).
3. Wunsch, C. & Ferrari, R. *Annu. Rev. Fluid Mech.* **36,** 281–314 (2004).
4. Munk, W. & Wunsch, C. *Deep Sea Res. I* **45,** 1977–2010 (1998).
5. Garrett, C. & Kunze, E. *Annu. Rev. Mar. Sci.* **39,** 57–87 (2007).
6. Nikurashin, M. & Ferrari, R. *Geophys. Res. Lett.* **38,** L08610 (2011).
7. Toggweiler, J. R. & Samuels, B. *J. Phys. Oceanogr.* **28,** 1832–1852 (1998).
8. Ferrari, R. *et al. Proc. Natl Acad. Sci. USA* **111,** 8753–8758 (2014).
9. Lumpkin, R. & Speer, K. *J. Phys. Oceanogr.* **37,** 2550–2562 (2007).

SENSORY SYSTEMS

Sound processing takes motor control

Neurons linking the brain region that controls movement to the region involved in auditory control have been found to suppress auditory responses when mice move, but the reason for this inhibition is unclear. See Article p.189

URI LIVNEH & ANTHONY ZADOR

The key to human cognition lies in the neocortex, a modular brain structure that is unique to mammals. Within each neocortical module, small ensembles of neurons are wired together in stereotyped patterns. Subsets of these neurons send long-range axonal projections to other modules to create systems of circuits that transform the activity of single neurons into complex behaviours such as perception, cognition and motor control. Understanding how different neocortical regions — including the motor, visual and auditory cortices — coordinate their activity is a central challenge in systems neuroscience. In this issue, Schneider *et al.*[1] (page 189) describe a technically sophisticated set of experiments that unravels the mechanisms by which the motor cortex exerts control over the auditory cortex during locomotion.

Locomotion facilitates visual responses in the visual cortex[2] but, conversely, Schneider and colleagues observed that it suppresses sound-evoked responses in the auditory cortex. This observation is intriguing because these responses are also suppressed when an animal vocalizes[3] or engages in an auditory task[4], behavioural states that require careful auditory processing. What is the mechanism by which locomotion suppresses neuronal responses in the auditory cortex?

Neuronal firing rates are determined by the balance between signals that promote and inhibit firing, so, in principle, firing can be suppressed by either a decrease in excitatory signals or increased inhibition. To distinguish between these possibilities, Schneider and co-workers performed the challenging feat of making intracellular-activity recordings from neurons in the auditory cortex of mice running on a treadmill. These experiments revealed that decreased auditory responses during locomotion are the result of an increase in inhibition. Cortical inhibition arises almost entirely from local inhibitory interneurons that make only short-range connections with nearby neurons, so the interneurons are probably driven by long-range excitatory inputs that transmit signals into the auditory

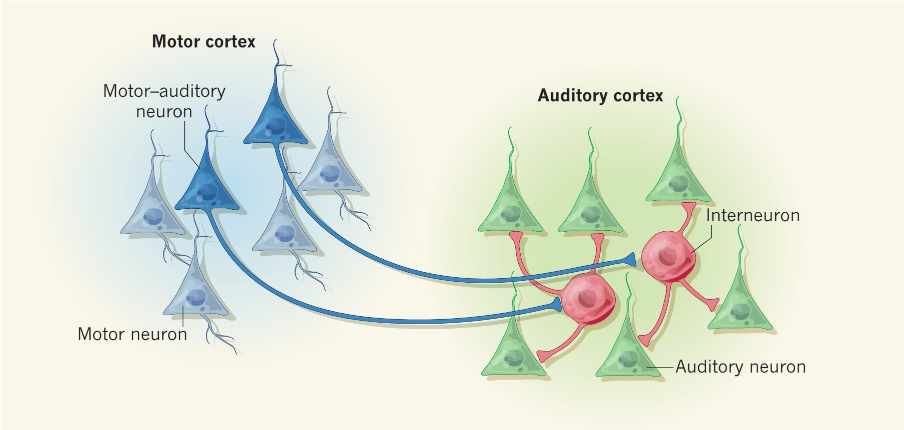


Figure 1 | Quiet in the auditory cortex. Schneider *et al.*[1] report that responses in the auditory cortex of the brain are suppressed during locomotion. When mice move, a subset of neurons in the motor cortex (motor–auditory neurons) sends excitatory signals to the interneurons of the auditory cortex, which in turn inhibit auditory neurons.

cortex. But which long-range inputs are responsible?

The authors hypothesized that long-range inputs arrive from the motor cortex. To test this, they labelled the subset of motor-cortex neurons that sends axonal projections to the auditory cortex (motor–auditory neurons) with a protein that fluoresces when activated, and monitored the neurons during locomotion. They found that the activity of motor–auditory neurons is increased before and throughout movement, indicating that they could be responsible for auditory-cortex suppression (Fig. 1). The researchers therefore set out to demonstrate that activation of motor–auditory neurons was not just correlated with suppression, but was also causally involved.

To establish causality, Schneider *et al.* infected motor–auditory neurons with a virus that enabled them to express channelrhodopsin-2 protein. Expression of channelrhodopsin-2 (which is originally derived from algae[5]) allows neurons to be activated in response to light. Selective stimulation of the axon terminals of motor–auditory neurons with light resulted in a suppression of the auditory cortex that was indistinguishable from that elicited by locomotion, supporting a causal role for this direct projection. However, this experiment alone was inconclusive, because excitation of motor–auditory axons may travel backwards along the motor projection, exciting other targets of the motor neurons and so indirectly affecting auditory responses. To rule out the possibility that suppression was indirect, the authors repeated the experiments while pharmacologically blocking activity in the motor cortex, and achieved the same result.

Finally, Schneider and colleagues inhibited motor-cortex neurons during locomotion, which disabled motor inputs to the auditory cortex. The authors found that in the absence of motor-cortex activity, locomotion was not accompanied by auditory suppression. Thus, the motor-to-auditory cortex projection is both necessary and sufficient for locomotion to suppress auditory responses.

Why should the auditory cortex be suppressed during locomotion? One might imagine that decreased activity in the auditory cortex implies reduced auditory sensitivity. However, behavioural conditions that require enhanced auditory processing typically suppress responses in the auditory cortex[3,4], raising the possibility that suppressed responsiveness serves to increase sensitivity. Such a seemingly paradoxical increase in sensitivity in the face of a general decrease in auditory cortical activity would occur if a privileged subset of cortical outputs were spared the effect of feedback suppression. In much the same way that shushing a noisy audience makes it possible to hear the seminar speaker, so feedback suppression may act to 'shush' all but the most important outputs from the auditory cortex.

The current results might be best considered in the framework of active sensation — that is, how animals separate self-induced sensory inputs from externally induced ones[6]. Movement and locomotion generate various types of self-induced sensation (for example, the movement of an object on your retina as you move your head), and so sensory inputs consist of both externally derived and self-induced sensations. Our perception separates these two sources of sensation to provide us with a movement-independent representation of the environment. To achieve this separation, a copy of the motor command might be used to indicate to the sensory cortices that movement is occurring. This copy could then be used to subtract the self-induced motor signal from the externally generated signal. The present results provide a detailed description of a circuit that may be involved in just such a computation. ■

Uri Livneh *and* **Anthony Zador** *are at Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA.*
e-mail: zador@cshl.edu

1. Schneider, D. M., Nelson, A. & Mooney, R. *Nature* **513,** 189–194 (2014).
2. Fu, Y. *et al. Cell* **156,** 1139–1152 (2014).
3. Eliades, S. J. & Wang, X. *Nature* **453,** 1102–1106 (2008).
4. Otazu, G. H., Tai, L.-H., Yang, Y. & Zador, A. M. *Nature Neurosci.* **12,** 646–654 (2009).
5. Yizhar, O., Fenno, L. E., Davidson, T. J., Mogri, M. & Deisseroth, K. *Neuron* **71,** 9–34 (2011).
6. Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H. & Lakatos, P. *Curr. Opin. Neurobiol.* **20,** 172–176 (2010).

This article was published online on 27 August 2014.

ASTROPHYSICS

Quasar complexity simplified

An analysis of a sample comprising some 20,000 mass-accreting supermassive black holes, known as quasars, shows that most of the diverse properties of these cosmic beacons are explained by only two quantities. **See Letter p.210**

MICHAEL S. BROTHERTON

If a picture is worth a thousand words, then a spectrum can be worth a thousand pictures. That is perhaps an underestimate when dealing with star-like blobs of light that look fuzzy even through the world's largest telescopes, as is the case with quasars. First recognized more than five decades ago as counterparts to radio sources[1], these extremely energetic entities are supermassive black holes in the nuclei of distant galaxies[2]. The black holes themselves do not emit light, but their gravity accelerates gas into swirling accretion disks that can outshine the galaxies they dwell in. Determining the physical properties of these systems from spectroscopic observations is challenging. But a study by

NRAO/AUI

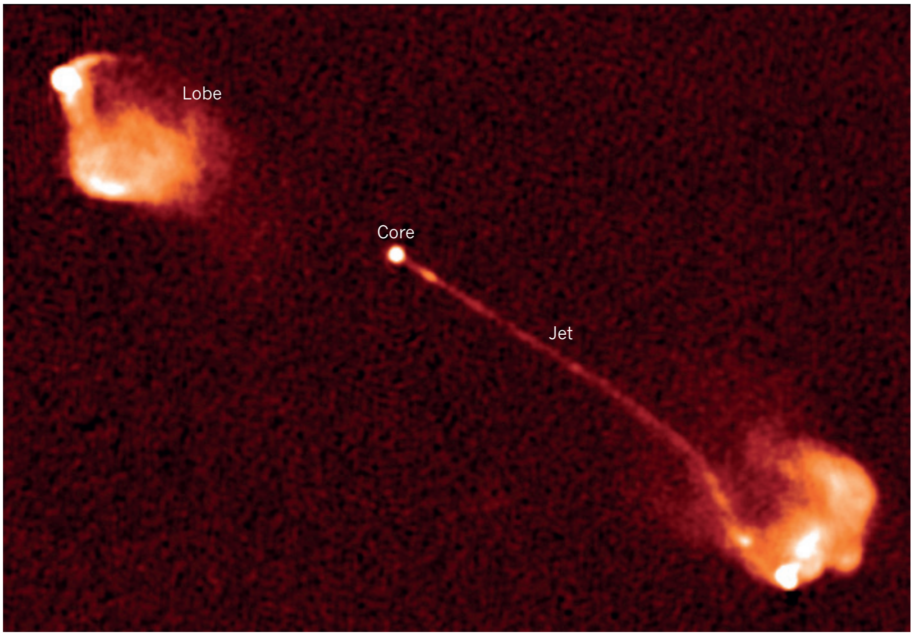


Figure 1 | Helpful jets. Quasar 3C 175 shows a prominent core, jet and lobes when mapped at radio wavelengths. An accretion disk, which feeds a supermassive black hole at the centre of the quasar's host galaxy, powers jets that extend far into intergalactic space. These large-scale structures provide a means of estimating the orientation of a quasar's axis of symmetry. Only a small percentage of quasars show such strong, clear radio jets as those pictured. Shen and Ho[3] have developed a method for determining the orientation of more-typical quasars on the basis of their optical spectra alone.

Shen and Ho[3] reported on page 210 of this issue accomplishes that elusive feat in the clearest way so far, using not one quasar spectrum, but the spectra of more than 20,000 quasars.

Visible light from a quasar has two main sources: a continuous spectrum emitted by the hot accretion disk, and discrete line emission from gas clouds that orbit the black hole and disk and that are ionized by the disk's intense radiation. The emission lines reveal information about the local environment. In particular, intensity ratios of emission lines due to different levels of ionization depend on the characteristics of the disk's radiation field. Many of these spectral properties are correlated in systematic ways[4] (a set of correlations referred to as 'Eigenvector 1', or EV1), which suggests that they are driven by one fundamental physical parameter: the Eddington ratio. This is the ratio of the luminosity of the quasar to that of a black hole that has maximal gas accretion — a limit reached when the quasar's radiation pressure balances its gravity.

Another major property of the emission lines is their width, measured as the full-width at their half-maximum (FWHM) intensity. The lines are broadened by the Doppler effect: gas in the line-emitting clouds moving away from Earth emits light shifted to longer wavelengths (redshifted), and gas moving towards us emits light shifted to shorter wavelengths (blueshifted). The FWHM of broad lines, in particular that of the Hβ hydrogen line ($FWHM_{H\beta}$), provides the component along Earth's line of sight of the gas's orbital velocity. The gravitational field associated with supermassive black holes, which are millions to billions of times more massive than the Sun, sets those velocities, and so measurements of $FWHM_{H\beta}$ help to determine a black hole's mass.

However, measurements of $FWHM_{H\beta}$ depend on how the quasar is tilted relative to our perspective. The large-scale jets that quasars emit, and which are seen in radio observations (Fig. 1), permit a determination of the jet's orientation, together with the quasar's symmetry axis, around which the accretion disk and the line-emitting gas clouds rotate. The value of $FWHM_{H\beta}$, and the inferred velocity, tend to be small when the jets are pointing in our direction, and large when they are pointing away[5]. But we do not clearly see strong, distinct jets in the general quasar population. Therefore, correcting for the geometric effect of the quasar tilt on $FWHM_{H\beta}$-based estimations of black-hole masses has not generally been possible.

Shen and Ho have tackled this problem. Making new and convincing arguments, they have plotted what they describe as "a main sequence of quasars" (see Fig. 1 of the paper[3]). The horizontal axis is the emission-line intensity ratio of iron (Fe II) to Hβ, denoted $R_{\mathrm{Fe\,II}}$, which characterizes EV1 and also tracks the Eddington ratio; the vertical axis is $FWHM_{H\beta}$, which segregates quasars for a given EV1 by the orientations of the systems. Shen and Ho's Figure 1 allows astronomers to go from spectral properties that are easily measured — $FWHM_{H\beta}$ and $R_{\mathrm{Fe\,II}}$ — to two fundamental quantities that account for the observed diversity of quasars: the Eddington ratio and orientation.

Furthermore, the authors have, for the first time, reported a statistically significant difference in the large-scale environments of quasars with differing Eddington ratios. They find that the more-massive black holes, which have lower accretion rates and hence lower Eddington ratios, exist in large-scale environments in which quasars and their host galaxies are more strongly clustered — in accordance with theoretical expectations. Tying the properties of quasars, which operate on tiny scales compared with the galaxies that harbour them, to even grander large-scale structure is a most intriguing development. The behaviour of galactic nuclei is thus linked to the largest scales of galaxy clusters, indicating evolutionary relationships between these two entities on cosmic scales.

A century ago, stellar astronomy underwent a similar breakthrough in linking observable parameters to more-fundamental physical quantities. By plotting the colours of stars against their luminosities, astronomers noticed a band, called the main sequence, along which most stars fall. The position of a star on this band is determined by its mass, which in turn governs many stellar properties, from temperature to size to lifetime. Quasars are very different from stars, as is their newly identified main sequence — which is perhaps more accurately described as a main wedge. But in an analogous way, we may now hope to develop a deeper understanding of quasars, their physical properties and perhaps even more. Clearly, the main sequence of quasars needs further testing, and only time will tell whether its utility is equal to that of the stellar main sequence.

In any event, we are now better placed to use our telescopes for collecting spectra and other data from quasars, and in turn to associate a set of observed data with a set of physical parameters. The quasar research field is maturing, just as stellar astronomy once did. Even though a quasar's Eddington ratio, black-hole mass, inclination angle, luminosity and other properties all simultaneously affect its spectrum, establishing a main sequence promises to provide an invaluable tool for separating out competing effects. Perhaps some day, advances in technology will allow us to obtain an image of a quasar that is clear enough to verify the picture that Shen and Ho advance today. ■

Michael S. Brotherton *is in the Department of Physics & Astronomy, University of Wyoming, Laramie, Wyoming 82071, USA.*
e-mail: mbrother@uwyo.edu

1. Schmidt, M. *Nature* **197,** 1040 (1963).
2. Lynden-Bell, D. *Nature* **223,** 690–694 (1969).
3. Shen, Y. & Ho, L. C. *Nature* **513,** 210–213 (2014).
4. Boroson, T. A. & Green, R. F. *Astrophys. J. Suppl. Ser.* **80,** 109–135 (1992).
5. Wills, B. J. & Browne, I. W. A. *Astrophys. J.* **302,** 56–63 (1986).

ARTICLE

doi:10.1038/nature13711

Assembly-line synthesis of organic molecules with tailored shapes

Matthew Burns[1], Stéphanie Essafi[1], Jessica R. Bame[1], Stephanie P. Bull[1], Matthew P. Webster[1], Sébastien Balieu[1], James W. Dale[2], Craig P. Butts[1], Jeremy N. Harvey[1] & Varinder K. Aggarwal[1]

Molecular 'assembly lines', in which organic molecules undergo iterative processes such as chain elongation and functional group manipulation, are found in many natural systems, including polyketide biosynthesis. Here we report the creation of such an assembly line using the iterative, reagent-controlled homologation of a boronic ester. This process relies on the reactivity of α-lithioethyl tri-isopropylbenzoate, which inserts into carbon–boron bonds with exceptionally high fidelity and stereocontrol; each chain-extension step generates a new boronic ester, which is immediately ready for further homologation. We used this method to generate organic molecules that contain ten contiguous, stereochemically defined methyl groups. Several stereoisomers were synthesized and shown to adopt different shapes—helical or linear—depending on the stereochemistry of the methyl groups. This work should facilitate the rational design of molecules with predictable shapes, which could have an impact in areas of molecular sciences in which bespoke molecules are required.

Nature has evolved highly sophisticated machinery for organic synthesis. An archetypal example is its machinery for polyketide synthesis where a simple thioester is passed from one module to another, undergoing enzyme-catalysed acylation, dehydration, reduction or chain extension multiple times until the target molecule is formed[1]. The process amounts to a molecular assembly line. By iteration and variation of the processing enzymes, nature manufactures an enormously diverse array of polyketides, many of which display high chemical complexity and biological activity (Fig. 1a).

We have sought to emulate nature in a related approach, but using boronic esters rather than thioesters. Our approach is to develop reagents which insert into the C–B bond, and to carry out this process iteratively so that a simple boronic ester is ultimately converted into a complex molecule with full control over its length, its shape and its functionality (Fig. 1b). By making specific molecules in this way, we also planned to obtain further understanding of the role of methyl substituents that are often interspersed along flexible carbon chains in natural products. The methyl groups originate from the metabolism of propionates or by methylation reactions, but nature could equally well have used acetates instead or avoided methylation and so managed with less complex machinery and created less complex organic molecules. The seemingly trivial substitution of a hydrogen atom for a methyl group on a carbon chain must have a powerful underlying evolutionary advantage. It has been suggested that nature uses the methyl groups (together with other polar residues) to give the molecule a predisposition to adopt the required conformation for interaction with its biological target without significant loss of enthalpy or entropy[2–8]. Despite this structural predisposition, the molecule is still flexible enough to change its shape when required (for example for transport across membranes). To probe the singular effect of how methyl substituents affect conformation of carbon chains it would be desirable to make molecules with multiple contiguous methyl groups, but such molecules were previously deemed impossible to prepare[2], in contrast to 1,3-deoxypolypropionates[9,10].

In this Article, we report our success in making such molecules through a highly streamlined process. In particular, through an iterative homologation procedure, we have developed a highly selective assembly-line synthesis process, and by successfully targeting carbon chains carrying ten contiguous methyl groups and no other functionality, we show that the methyl substituents can be used to control the conformation of the molecule with great precision. They act as levers, pushing or pulling the carbon chain and, depending on their specific orientation, they can force the molecule to adopt a linear or helical conformation both in solution and in the solid state. This is analogous to the way in which the primary sequence of amino acids determines their folded shape[11]. Indeed, the iterative homologation process we have developed makes it possible to rationally design and create molecules with predictable shapes without having to incorporate functional groups to bias a particular conformation.

Development of the iterative homologation process

Two broad approaches have been developed for the stereocontrolled homologation of boronic esters: a substrate-controlled method in which a chiral diol on the boronic ester controls the stereochemistry (Matteson homologation[12,13]), and a reagent-controlled method in which chirality in the reagent controls the stereochemistry[14]. The latter method is more direct and more versatile because it enables ready access to alternative stereoisomers. We (and others[15–18]) have focused on reagent-controlled methodology. We have found that Hoppe's lithiated carbamates[19–21] homologate boronic esters with high stereocontrol, and we have applied this methodology in the synthesis of a number of natural products[22–26]. To apply this methodology to iterative homologations, we set the goal of creating a carbon chain with ten contiguous methyl substituents with total stereocontrol (Fig. 1b). This is a daunting task because each step must go to completion without over- or under-homologation and it must proceed with full stereocontrol to obtain pure material, because different chain lengths and different diastereoisomers would be extremely difficult to separate. As an illustration, if each homologation occurred in 98% completion with just 1% over-homologation and 1% under-homologation, then after ten iterative homologations a binomial distribution of products would be obtained in which the major compound, having undergone 10 homologations (a 10-mer), was only 82% pure, contaminated by under- and over-homologation products. If each homologation occurred with 98:2 enantioselectivity, then after ten iterative homologations the product would be a mixture of diastereoisomers that was only 82% pure, which again was undesirable. As a further illustration of the challenge,

[1]School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK. [2]Novartis Horsham Research Centre, Wimblehurst Road, Horsham, West Sussex RH13 5AB, UK.

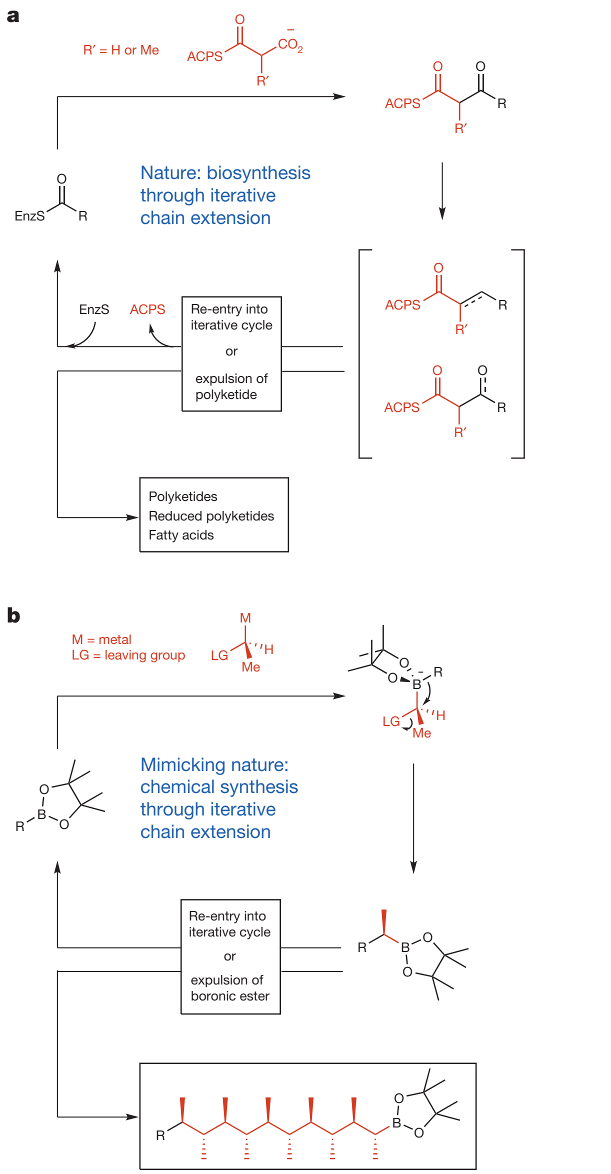


Figure 1 | Iterative approaches to assembly-line synthesis. **a**, Example of polyketide biosynthesis where successive cycles of chain extension and functional-group interconversions generate a diverse array of complex molecules. ACP, acyl carrier protein; Enz, enzyme; Me, methyl. **b**, Proposed reagent-controlled homologation of boronic esters where successive cycles of chain extension enable rapid and streamlined synthesis of stereodefined carbon chains.

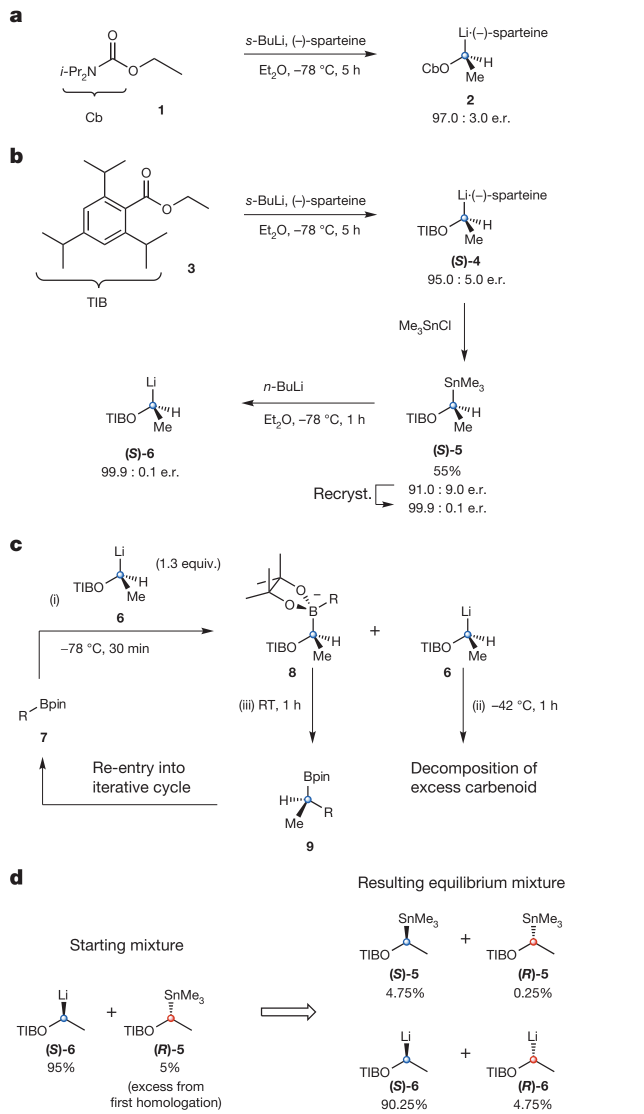


Figure 2 | Methodology used for homologation of boronic esters. **a**, Method for the generation of Hoppe's lithiated carbamate. Bu, butyl; Et, ethyl; *i*-Pr, isopropyl; OCb, *N*,*N*-diisopropyl carbamate. **b**, Method for the generation of α-lithiated hindered benzoate **6** with high enantiomeric ratio from stannane **5**. TIB, 2,4,6-triisopropylbenzoate. **c**, Optimized protocol for iterative homologation of boronic esters. Carbenoid **2** was not suitable for iterative homologations whereas carbenoid **6** was suitable and the protocol for its successful use is shown here. pin, pinacol. RT, room temperature. **d**, Racemization pathway for lithiated benzoate (***S***)-**6** when an excess of stannane (***R***)-**5** is present from the previous homologation. This example shows the ratio of products obtained from a mixture of lithiated benzoate (***S***)-**6** (95%, 99.9:0.1 e.r.) and stannane (***R***)-**5** (5%, 99.9:0.1 e.r.), which leads to lithiated benzoate and stannane of lower enantiomeric ratio (~95:5).

a recent triple, one-pot homologation of a boronic ester using an α-chloroalkyllithium reagent yielded[27], in the best case, a mixture of the trimer (19%, 5:80:9:6 diastereomeric ratio), dimer (27%, 78:22 diastereomeric ratio) and monomer (5%).

Unfortunately, Hoppe's lithiated carbamates (**2**) (Fig. 2a), which we had used extensively in synthesis, could not be employed because we found that they were prone to giving significant quantities of over- and under-homologated products with hindered boronic esters, and that furthermore they could be obtained in only 97:3 enantiomeric ratio (e.r.). We therefore turned to α-metallated hindered benzoates as alternatives to Hoppe's carbamates, because we had found that the superior leaving-group ability of the ester relative to the carbamate enabled difficult homologations to proceed more effectively[28]. After exploring various alternatives, we found that deprotonation of ethyl tri-isopropylbenzoate **3** with *s*-BuLi/(−)-sparteine[28] followed by trapping with Me_3SnCl gave stannane **5** (91:9 e.r.), which could be recrystallized to 99.9:0.1 e.r. The required enantioenriched lithiated benzoate **6** could be easily generated from stannane **5** with *n*-BuLi with retention of stereochemistry[29] (Fig. 2b). Using this method, both enantiomers of the stannane were easily prepared on multigram scale, using commercially available (+)- or (−)-sparteine, without the need for chromatography. In addition, the chiral diamines were re-isolated (and reused) in >80% yield. Having access to substantial quantities of both enantiomers of these derivatives greatly facilitated the iterative homologation process; it was like having the chiral organometallic **6** in a bottle[30].

An optimized protocol had to be developed for iterative homologation to ensure high fidelity (Fig. 2c). Treatment of stannane **5** with *n*-BuLi at −78 °C followed by addition of a boronic ester **7** gave the boronate complex **8**. An excess of stannane **5** (and, therefore, an excess of lithiated benzoate **6**) was used to ensure that all of the boronic ester

was converted into the boronate complex. The reaction mixture was then kept at −42 °C for one hour, to allow excess lithiated benzoate **6** to decompose (Supplementary Information). At this temperature, the boronate complex is stable, but, after warming to room temperature (20 °C) for one hour, the 1,2-migration occurred, giving the homologated product **9**. The ageing at −42 °C for one hour is essential to prevent a small amount (about 0.5%) of the product boronic ester reacting with the excess lithiated benzoate and giving over-homologated product. The reaction was then filtered to remove the insoluble lithium salt of 2,4,6-triisopropylbenzoic acid, to give the crude boronic ester that was used directly in a subsequent homologation. Although we were able to conduct up to seven homologations iteratively (without aqueous work-up or purification of any intermediates) and obtain pure material, we found it more reliable to remove by-products using an aqueous work-up after every third homologation.

Having developed an optimized protocol for the homologation of boronic esters, we set about conducting the iterative homologation sequence. We initiated our sequence from biphenyl boronic ester **(*R*)-10** rather than from 4-biphenylboronic acid pinacol ester owing to the latter's poor solubility in Et_2O. Boronic ester **(*R*)-10** was subjected to nine consecutive homologations using lithiated benzoate **(*S*)-6**, with an aqueous work-up being performed after every third homologation, giving boronic ester **11** in 58% yield (Fig. 3a). Each homologation was followed by gas chromatography mass spectrometry (GC–MS), which indicated that very low levels of over- and under-homologation occurred. In fact, at the end of the sequence the product was a 1:97:2 ratio of 9-mer:10-mer:11-mer, demonstrating the extraordinarily high fidelity of each homologation reaction. ^{1}H and ^{13}C NMR showed that it was also essentially one diastereoisomer. As a result of chiral amplification[31], it would also be a single enantiomer. On the basis of the Horeau principle[32], after nine homologations using stannane **5** (10^3:1 e.r.) on boronic ester **10** (~10^2:1 e.r.) the enantiomeric ratio of the major diastereoisomer should be 10^{29}:1, which is considerably greater than Avogadro's number, and so the product is expected to be literally a single enantiomer. An X-ray crystal structure of benzoate ester **12** confirmed the relative stereochemistry of the product.

Having demonstrated a highly effective assembly-line synthesis protocol, we sought to target other specific diastereoisomers. The conformation

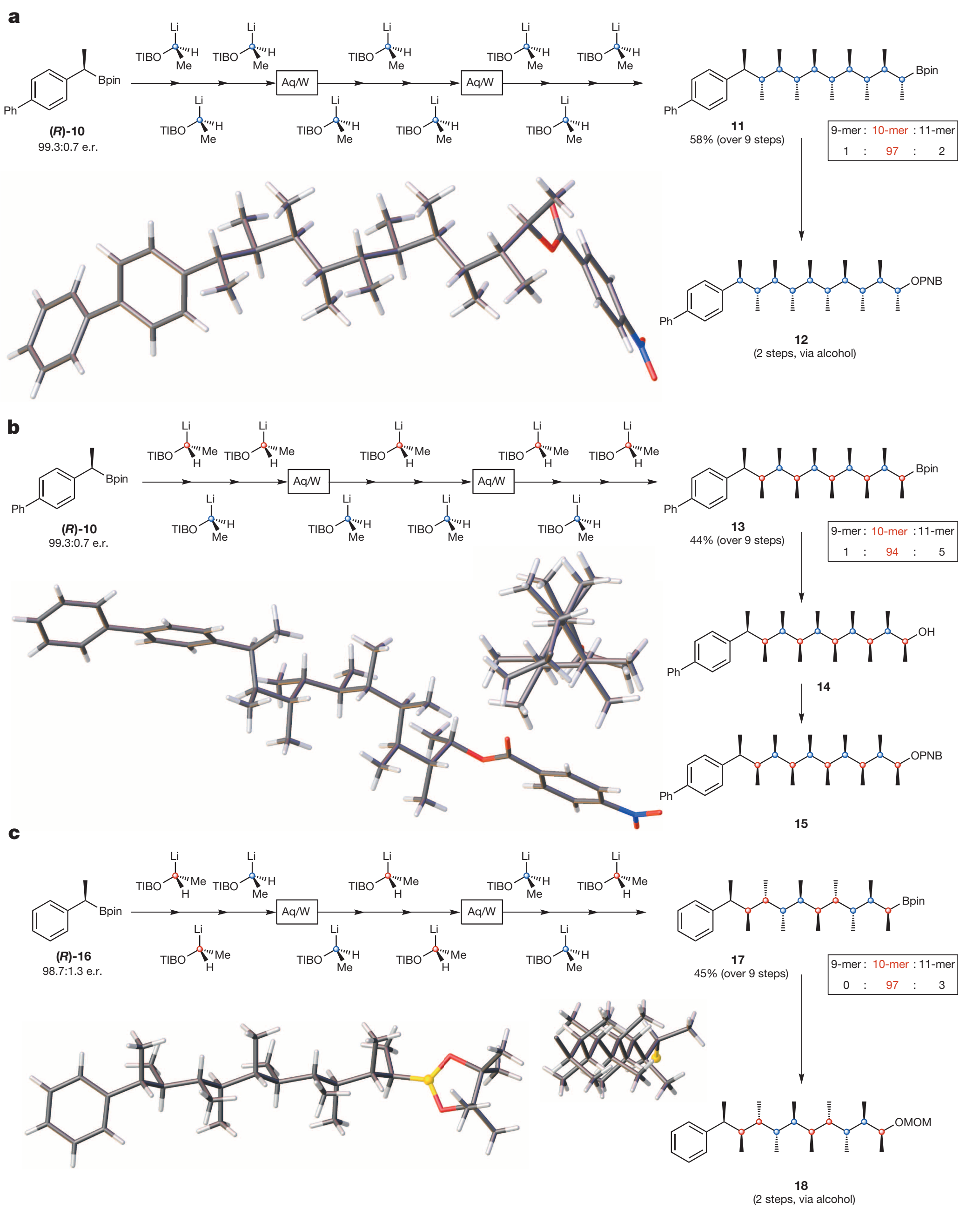


Figure 3 | Iterative assembly-line synthesis. a, Synthesis of the all-*anti* isomer boronic ester **11** and X-ray structure of the *p*-nitro benzoate derivative **12**. **b,** Synthesis of the all-*syn* isomer boronic ester **13** and X-ray structure (two views) of the *p*-nitro benzoate derivative **15**. **c,** Synthesis of the alternating *syn-anti* isomer boronic ester **17** with X-ray structure (two views) and the methoxymethyl ether (MOM) derivative **18**. The X-ray structures show that the all-*syn* isomer adopts a helical conformation, the alternating *syn-anti* isomer adopts a linear conformation, and the all-*anti* isomer does not adopt a regular conformation. Conditions for homologation: (1) addition of boronic ester to lithiated benzoate, −78 °C, 30 min; (2) −42 °C, 1 h; (3) room temperature, 1 h; (4) filter; (5) repeat. The ratios of boronic ester homologues were obtained by GC–MS analysis (Supplementary Information). Aq/W, aqueous work-up; Ph, phenyl; PNB, *p*-nitro benzoate. Blue and red dots specify which carbenoid (**6**) was used and incorporated into the product: blue dots represent **(*S*)-6** carbenoid derived from (−)-sparteine; red dots represent **(*R*)-6** carbenoid derived from (+)-sparteine.

of carbon chains should be controlled by *syn*-pentane interactions (also known as g^+g^- interactions) between the methyl groups[2,33] (Fig. 4a). Although the all-*anti* diastereoisomer **11** or **12** was not expected to adopt a particular low-energy conformation (as confirmed by the X-ray structure), we reasoned that the all-*syn* isomer **13** should adopt a helical conformation and the alternating *syn-anti* diastereoisomer **17** should adopt a linear conformation[2] (Fig. 4b, c). Our unique methodology provided a method to make such molecules and an opportunity to test whether *syn*-pentane interactions alone could control the chain conformation of otherwise flexible molecules.

Attempts to make the all-*syn* isomer, which required alternating between the enantiomers of the stannane, initially proved problematic. Analysis of the second homologation product showed that it was only a 95:5 mixture of diastereoisomers. Careful experimentation revealed the source of the problem. In the first homologation, stannane **(*R*)-5** was used in a slight excess (0.05 equiv. excess) over *n*-BuLi to ensure that no *n*-BuLi, which might react irreversibly with the boronic ester, remained. However, in the second homologation the slight excess of stannane **(*R*)-5** must have equilibrated with **(*S*)-6**, resulting in the reagent having a lower enantiomeric ratio, and so generated mixtures of diastereoisomers (Fig. 2d). The solution to the problem was to control the stoichiometry more precisely, that is, to use a 1.00:1.00 ratio of stannane **5** to *n*-BuLi. Using this modification, the assembly-line synthesis process was launched as before, alternating between the enantiomers of the stannane, and the all-*syn* isomer **13** was prepared in 44% yield (Fig. 3b). As before, each homologation was followed by GC–MS, and at the end of the sequence the product was a 1:94:5 ratio of 9-mer:10-mer:11-mer, again demonstrating the extraordinarily high fidelity of each homologation reaction. ^{1}H and ^{13}C NMR showed that it was also essentially one diastereoisomer and was expected to be a single enantiomer. An X-ray crystal structure of benzoate ester **15** confirmed the relative stereochemistry of the product and showed that in the solid state the flexible carbon backbone of the molecule adopted a perfect right-handed (*P*) helical conformation. The carbon chain does one complete turn every six carbon atoms in the backbone of the molecule.

The synthesis of the alternating *syn-anti* diastereoisomer required the use of alternating pairs of the stannane enantiomers. Performing the iterative homologation from biphenyl boronic ester **(*R*)-10** led to insoluble intermediates after six homologations, and so the phenyl analogue **(*R*)-16** was used instead. Re-launching the iterative homologation sequence as before from boronic ester **(*R*)-16**, with an aqueous work-up after every third homologation, gave boronic ester **17** in 45% yield (Fig. 3c). As before, each homologation was followed by GC–MS, and at the end of the sequence the product was a 0:97:3 ratio of 9-mer:10-mer:11-mer. ^{1}H and ^{13}C NMR showed that it was also essentially one diastereoisomer and was expected to be a single enantiomer. The boronic ester **17** was itself crystalline, and X-ray crystallography not only confirmed its structure but also showed that the molecule adopted a perfectly linear conformation.

NMR and computational analysis of solution structures

The X-ray structures show that in the solid state, the all-*syn* isomer **15** adopts a helical conformation and that the alternating *syn-anti* isomer **17** adopts a linear conformation, as predicted, on the basis of conformational control dominated by minimizing *syn*-pentane interactions. This is reminiscent of the difference between the respective structures adopted by isotactic and syndiotactic polypropylene[34] and O'Hagan's polyfluorinated alkanes[35]. In the case of polypropylene the isotactic form is helical, whereas the syndiotactic form exhibits more complex behaviour, with linear structures in some cases. In the case of polyfluorinated alkanes, the all-*syn* isomer adopted a helical conformation and the alternating *syn-anti* isomer adopted a linear conformation, although here the preference was dominated by electronic rather than steric effects. However, crystal packing will influence conformation in the solid state, and we were keen to examine whether similar conformations existed in solution. We note that although small molecules with only one or two rotatable bonds can adopt predominantly one conformation, larger molecules with multiple rotatable bonds do not, because the enthalpic gain in minimizing *syn*-pentane interactions is outweighed by the entropic cost of greater rigidity: this is illustrated by alkanes **19**, **20** and **21**, which are of increasing chain length and showed 91%, 76% and just 58% preference for a single conformer[36] (Fig. 4d). It is therefore extremely difficult to make larger molecules that adopt largely one conformation. In our case, we expected an increase in the enthalpic gain in minimizing *syn*-pentane interactions because we have twice the number of such interactions with no increase in the number of rotatable bonds, and so expected a higher preference for a single conformer. However, this gain is moderated by the increase in the number of gauche interactions, which reduces the difference in energy between different conformers. We therefore embarked on establishing the solution conformation using a combined experimental and computational approach.

Despite the very congested nature of the spectra, NMR spectroscopy of all-*syn* isomer **14** and *syn-anti* isomer **18** yields an extensive set of accurate interproton distances[37,38] (derived from measurements of the nuclear Overhauser effect) and ^{1}H–^{1}H and ^{1}H–^{13}C couplings. Using *ab initio* calculated structures, relative free energies and spin–spin couplings for the family of conformers for each species, these observations can be deconvolved to provide an integrated overview of solution behaviour (Fig. 5).

Molecular mechanics was used to exhaustively explore conformational space for model compounds **14a** and **18a** (analogues of **14** and **18**, respectively, but with truncated end-groups; 9,710 and 3,970 conformers were respectively assessed). Refined structures, population estimates and predicted *J* couplings were then obtained using electronic structure theory (MP2 with explicit correlation[39] using density functional theory-optimized structures and a continuum solvation correction) for a smaller number of conformers of **14a** and **18a** (66 and 50, respectively), on the basis of all low-energy molecular mechanics structures plus a few manually selected structures suggested by the NMR analysis.

In both cases, calculations predict that several different conformers are populated at room temperature; however, in each case, the preference for linear and helical structures is very strong (Fig. 5a) with each C–C bond along the carbon chain dominated by a single dihedral angle. For **18a**, one linear conformer is predicted to represent 95% of the population, versus 74% for a helical conformer in the case of **14a** (ignoring end-group rotamers). This is one of the highest preferences found so far for a flexible molecule, and highlights the value in maximizing the density of *syn*-pentane interactions to control conformation. The calculated and measured NMR properties for **14**/**14a** and **18**/**18a** agree very well, provided that all reasonably highly populated conformers are included (Fig. 5b), and the levels of agreement are in line with those obtained for accurate conformational assignments in less flexible molecules[37,40,41]. The corresponding alcohol of the all-*anti* isomer boronic ester **11** appears highly disordered by NMR and computational analysis, as predicted (Supplementary Information).

The solution properties thus show that these systems behave as ensembles of structures from which the dominant linear and helical character clearly emerges. The calculated major low-energy conformers for the helical molecule **14a** and the linear molecule **18a** are shown in Fig. 5c. However, although a tight helix is observed in the X-ray structure of **15**, which completes one turn for every six carbon atoms, we found that the solution conformation of **14**/**14a** is a loose helix, where the carbon chain completes one turn for every nine carbon atoms. This loosening results from significant deviation of the dihedral angles along the chain from the ideal angles (60° or 180°) observed in the solid state. As shown in Fig. 4e, as the chain coils the greater interaction between the (larger) alkyl chains forces the chains apart slightly (>60° dihedral angles), resulting in a small opening of the helix that is counterbalanced by a decrease in the angle between the smaller methyl groups[33] (<60° dihedral angles).

In solution a left-handed helix (*M*) is observed for **14**/**14a**, whereas in the solid state the same diastereoisomer **15** adopts a right-handed helix (*P*). This is probably due to the different end-groups (OH versus

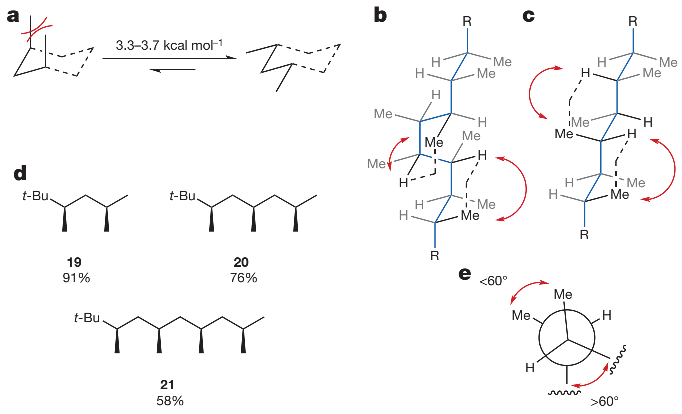


Figure 4 | The effect of *syn*-pentane and other intramolecular steric interactions on conformation of molecules. **a**, Energy penalty incurred with a *syn*-pentane interaction. **b**, Expected helical conformation of the all-*syn* isomer **13**, where methyl groups along the carbon chain avoid *syn*-pentane interactions (red arrows). **c**, Expected linear conformation of the alternating *syn-anti* isomer **17**, where methyl groups along the carbon chain avoid *syn*-pentane interactions (red arrows). **d**, Carbon chains bearing *syn*-1,3-dimethyl units, and the percentage occupancy of a single dominant conformation[36]. **e**, Minor distortion in the conformation of the carbon chain of the all-*syn* isomer **14** (helical molecule), as determined by NMR and computational analysis. Because of 1,4-steric interactions, the carbon chain is pushed farther apart causing significant deviation from the ideal dihedral angles.

O-(4-nitro)benzoate). For efficient crystal packing, the nitrobenzoate group will prefer an extended (*anti*) conformation at θ_I (Fig. 5), acting as the principal large inductor group, and this terminal stereocentre induces the sense of handedness down the carbon chain[36]. Owing to the pseudo-change in chirality of this stereocentre, the opposite helical sense is created down the carbon chain.

Conclusion

We have developed a practical method for the reagent-controlled homologation of a boronic ester, which can be conducted iteratively and with total stereocontrol. Because these reactions are totally dominated by reagent control, no matched and mismatched effects are observed, enabling different stereoisomers to be targeted and prepared with equal ease. In addition, each chain extension step generates a new boronic ester, ready and primed for further homologation without requiring extra manipulation, making the process considerably more rapid and streamlined than alternative iterative strategies, which usually require several functional group interconversions between chain extension steps[42–45]. However, iterative Suzuki–Miyaura cross-coupling[46–48], zirconium-catalysed asymmetric carboalumination reactions[49] and triple-aldol cascade reactions[50] represent notable exceptions where additional steps between iterations are not required.

Thus, using our iterative homologation sequence we have been able to convert simple boronic esters into complex molecules bearing ten contiguous methyl substituents with full stereocontrol. Different stereoisomers have been targeted and their conformations were determined by X-ray crystallography and NMR and analysed computationally. All three methods of analysis showed that both in the solid state and in solution the all-*anti* isomer did not adopt a particular conformation, whereas the all-*syn* isomer adopted a helical conformation and the alternating *syn-anti* isomer adopted a linear conformation. In the latter two cases, the methyl groups along the carbon chain were able to force the molecule to adopt these particular conformations as a result of *syn*-pentane interactions alone. By incorporating the effect of *syn*-pentane interactions on conformation and using the iterative homologation process we have developed, molecules with predictable shape can now be rationally designed and created. This should have an impact in all areas of molecular sciences where bespoke molecules are required.

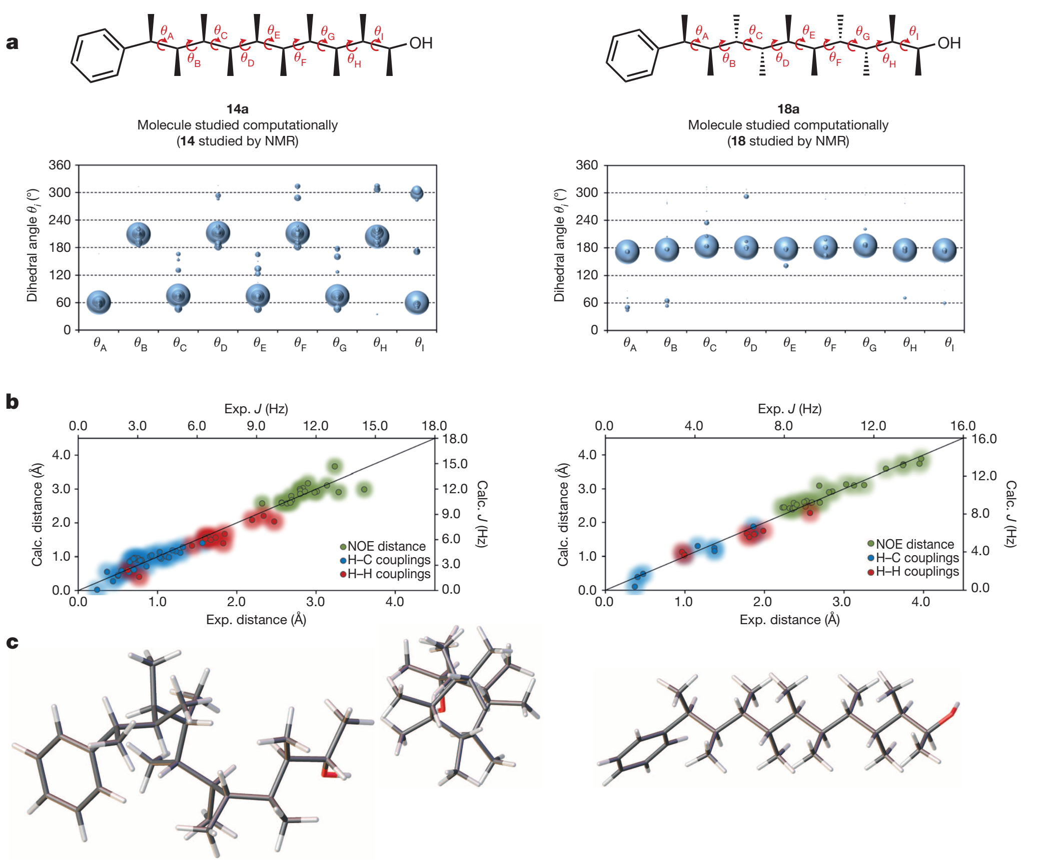


Figure 5 | Solution conformations of compounds 14 and 18. **a**, Theoretically predicted properties of the ensemble of conformations of model compounds **14a** and **18a**. Each populated conformer is shown as nine dots, of size proportional to the calculated relative abundance of that conformer, and with a position defined by the calculated value of the backbone dihedral angle, θ_i (i = A, …, I). **b**, Correlation between theoretically predicted NMR properties (interproton distance, green (NOE; nuclear Overhauser effect); ^{1}H–^{1}H and ^{1}H–^{13}C scalar J couplings, respectively red and blue) for the ensemble of conformers of **14a** (left) and **18a** (right) and experimentally observed values (interproton distance, bottom; scalar J couplings, top) of **14** and **18**. Each dot is associated with errors of ~5% in interproton distance and approximately ±1 Hz in J. The calculations predict **14a** to be predominantly helical in nature, with **18a** overwhelmingly populating linear conformers. NMR measurements in solution are completely in line with this predicted behaviour. **c**, Structure of the calculated dominant conformations of **14a** (left) and **18a** (right). Grey, carbon; white, hydrogen; red, oxygen.

Received 9 May; accepted 23 July 2014.

1. Staunton, J. & Weissman, K. J. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* **18,** 380–416 (2001).
2. Hoffmann, R. W. Flexible molecules with defined shape: conformational design. *Angew. Chem. Int. Edn Engl.* **31,** 1124–1134 (1992).
3. Hoffmann, R. W. Conformation design of open-chain compounds. *Angew. Chem. Int. Edn Engl.* **39,** 2054–2070 (2000).
4. Smith, P. W. & Still, W. C. The effect of substitution and stereochemistry on ion binding in the polyether ionophore monensin. *J. Am. Chem. Soc.* **110,** 7917–7919 (1988).
5. Wang, X., Erickson, S. D., Iimori, T. & Still, W. C. Enantioselective complexation of organic ammonium ions by simple tetracyclic podand ionophores. *J. Am. Chem. Soc.* **114,** 4128–4137 (1992).
6. Wei, A., Boy, K. M. & Kishi, Y. Biological evaluation of rationally modified analogs of the H-type II blood group trisaccharide. A correlation between solution conformation and binding affinity. *J. Am. Chem. Soc.* **117,** 9432–9436 (1995).
7. Boger, D. L., Ramsey, T. M., Cai, H., Hoehn, S. T. & Stubbe, J. Definition of the effect and role of the bleomycin A_2 valerate substituents: preorganization of a rigid, compact conformation implicated in sequence-selective DNA cleavage. *J. Am. Chem. Soc.* **120,** 9149–9158 (1998).
8. Nilewski, C., Geisser, R. W., Ebert, M.-O. & Carreira, E. M. Conformational and configurational analysis in the study and synthesis of chlorinated natural products. *J. Am. Chem. Soc.* **131,** 15866–15876 (2009).
9. Hanessian, S., Giroux, S. & Mascitti, V. The iterative synthesis of acyclic deoxypropionate units and their implication in polyketide-derived natural products. *Synthesis* **7,** 1057–1076 (2006).
10. ter Horst, B., Feringa, B. L. & Minnaard, A. J. Iterative strategies for the synthesis of deoxypropionates. *Chem. Commun.* **46,** 2535–2547 (2010).
11. Dill, K. A. & MacCallum, J. L. The protein-folding problem, 50 years on. *Science* **338,** 1042–1046 (2012).
12. Matteson, D. S. & Ray, R. Directed chiral synthesis with pinanediol boronic esters. *J. Am. Chem. Soc.* **102,** 7590–7591 (1980).
13. Matteson, D. S. Boronic esters in asymmetric synthesis. *J. Org. Chem.* **78,** 10009–10023 (2013).
14. Stymiest, J. L., Dutheuil, G., Mahmood, A. & Aggarwal, V. K. Lithiated carbamates: chiral carbenoids for iterative homologation of boranes and boronic esters. *Angew. Chem. Int. Ed.* **46,** 7491–7494 (2007).
15. Besong, G., Jarowicki, K., Kocienski, P. J., Sliwinski, E. & Boyle, F. T. Synthesis of (S)-(−)-*N*-acetylcolchinol using intramolecular biaryl oxidative coupling. *Org. Biomol. Chem.* **4,** 2193–2207 (2006).
16. Blakemore, P. R., Marsden, S. P. & Vater, H. D. Reagent controlled asymmetric homologation of boronic esters by enantioenriched main-group chiral carbenoids. *Org. Lett.* **8,** 773–776 (2006).
17. Blakemore, P. R. & Burge, M. S. Iterative stereospecific reagent-controlled homologation of pinacol boronates by enantioenriched α-chloroalkyllithium reagents. *J. Am. Chem. Soc.* **129,** 3068–3069 (2007).
18. Emerson, C. R., Zakharov, L. N. & Blakemore, P. R. Investigation of functionalized α-chloroalkyllithiums for a stereospecific reagent-controlled approach to the analgesic alkaloid (–)-epibatidine. *Chemistry* **19,** 16342–16356 (2013).
19. Hoppe, D., Hintze, F. & Tebben, P. Chiral lithium-1-oxyalkanides by asymmetric deprotonation; enantioselective synthesis of 2-hydroxyalkanoic acids and secondary alkanols. *Angew. Chem. Int. Edn Engl.* **29,** 1422–1424 (1990).
20. Hoppe, D. & Hense, T. Enantioselective synthesis with lithium/(−)-sparteine carbanion pairs. *Angew. Chem. Int. Edn Engl.* **36,** 2282–2316 (1997).
21. Beckmann, E., Desai, V. & Hoppe, D. Stereospecific reaction of α-carbamoyloxy-2-alkenylboronates and α-carbamoyloxy-alkylboronates with Grignard reagents - synthesis of highly enantioenriched secondary alcohols. *Synlett* **13,** 2275–2280 (2004).
22. Dutheuil, G., Webster, M. P., Worthington, P. A. & Aggarwal, V. K. Stereocontrolled synthesis of carbon chains bearing contiguous methyl groups by iterative boronic ester homologations: application to the total synthesis of (+)-faranal. *Angew. Chem. Int. Ed.* **48,** 6317–6319 (2009).
23. Robinson, A. & Aggarwal, V. K. Asymmetric total synthesis of solandelactone E: stereocontrolled synthesis of the 1,4-diol-2-ene core via lithiation-borylation-allylation sequence. *Angew. Chem. Int. Ed.* **49,** 6673–6675 (2010).
24. Pulis, A. P. & Aggarwal, V. K. Synthesis of enantioenriched tertiary boronic esters from secondary allylic carbamates. Application to the synthesis of C30 botryococcene. *J. Am. Chem. Soc.* **134,** 7570–7574 (2012).
25. Fletcher, C. J., Wheelhouse, K. M. P. & Aggarwal, V. K. Stereoselective total synthesis of (+)-giganin and its C10 epimer by using late-stage lithiation–borylation methodology. *Angew. Chem. Int. Ed.* **52,** 2503–2506 (2013).
26. Blair, D. J., Fletcher, C. J., Wheelhouse, K. M. P. & Aggarwal, V. K. Stereocontrolled synthesis of adjacent acyclic quaternary-tertiary motifs: application to a concise total synthesis of (–)-filiformin. *Angew. Chem. Int. Ed.* **53,** 5552–5555 (2014).
27. Sun, X. & Blakemore, P. R. Programmed synthesis of a contiguous stereotriad motif by triple stereospecific reagent-controlled homologation. *Org. Lett.* **15,** 4500–4503 (2013).
28. Larouche-Gauthier, R., Fletcher, C. J., Couto, I. & Aggarwal, V. K. Use of alkyl 2,4,6-triisopropylbenzoates in the asymmetric homologation of challenging boronic esters. *Chem. Commun. (Camb.)* **47,** 12592–12594 (2011).
29. Still, W. C. & Sreekumar, C. α-Alkoxyorganolithium reagents. A new class of configurationally stable carbanions for organic synthesis. *J. Am. Chem. Soc.* **102,** 1201–1202 (1980).
30. Rayner, P. J., O'Brien, P. & Horan, R. A. J. Preparation and reactions of enantiomerically pure α-functionalized Grignard reagents. *J. Am. Chem. Soc.* **135,** 8071–8077 (2013).
31. Negishi, E. A quarter of a century of explorations in organozirconium chemistry. *Dalton Trans.* 827–848 (2005).
32. Vigneron, J. P., Dhaenens, M. & Horeau, A. Nouvelle methode pour porter au maximum la purete optique d'un produit partiellement dedouble sans l'aide d'aucune substance chirale. *Tetrahedron* **29,** 1055–1059 (1973).
33. Tsuzuki, S. *et al.* Investigation of intramolecular interactions in *n*-alkanes. Cooperative energy increments associated with GG and GTG' sequences. *J. Am. Chem. Soc.* **113,** 4665–4671 (1991).
34. Lotz, B., Wittmann, J. C. & Lovinger, A. J. Structure and morphology of poly(propylenes): a molecular analysis. *Polymer* **37,** 4979–4992 (1996).
35. Hunter, L., Kirsch, P., Slawin, A. M. Z. & O'Hagan, D. Synthesis and structure of stereoisomeric multivicinal hexafluoroalkanes. *Angew. Chem. Int. Ed.* **48,** 5457–5460 (2009).
36. Hoffmann, R. W., Stahl, M., Schopfer, U. & Frenking, G. Conformation design of hydrocarbon backbones: a modular approach. *Chemistry* **4,** 559–566 (1998).
37. Butts, C. P., Jones, C. R. & Harvey, J. N. High precision NOEs as a probe for low level conformers – a second conformation of strychnine. *Chem. Commun.* **47,** 1193–1195 (2011).
38. Butts, C. P. *et al.* Interproton distance determinations by NOE – surprising accuracy and precision in a rigid organic molecule. *Org. Biomol. Chem.* **9,** 177–184 (2011).
39. Adler, T. B., Werner, H. J. & Manby, F. R. Local explicitly correlated second-order perturbation theory for the accurate treatment of large molecules. *J. Chem. Phys.* **130,** 054106 (2009).
40. Chini, M. G. *et al.* Quantitative ROE-derived interproton distances combined with quantum chemical calculations of NMR parameters in the stereochemical determination of conicasterol F, a nuclear receptor ligand from *Theonella swinhoei*. *J. Org. Chem.* **77,** 1489–1496 (2012).
41. Di Micco, S., Chini, M. G., Riccio, R. & Bifulco, G. Quantum mechanical calculation of NMR parameters in the stereostructural determination of natural products. *Eur. J. Org. Chem.* **2010,** 1411–1434 (2010).
42. Hanessian, S. *et al.* Application of conformation design in acyclic stereoselection: total synthesis of borrelidin as the crystalline benzene solvate. *J. Am. Chem. Soc.* **125,** 13784–13792 (2003).
43. Brand, G. J., Studte, C. & Breit, B. Iterative synthesis of (oligo)deoxypropionates via zinc-catalyzed enantiospecific sp^3-sp^3 cross-coupling. *Org. Lett.* **11,** 4668–4670 (2009).
44. ter Horst, B., Feringa, B. L. & Minnaard, A. J. Catalytic asymmetric synthesis of phthioceranic acid, a heptamethyl-branched acid from *Mycobacterium tuberculosis*. *Org. Lett.* **9,** 3013–3015 (2007).
45. Han, S. B., Hassan, A., Kim, I. S. & Krische, M. J. Total synthesis of (+)-roxaticin via C−C bond forming transfer hydrogenation: a departure from stoichiometric chiral reagents, auxiliaries, and premetalated nucleophiles in polyketide construction. *J. Am. Chem. Soc.* **132,** 15559–15561 (2010).
46. Lee, S. J., Gray, K. C., Paek, J. S. & Burke, M. D. Simple, efficient, and modular syntheses of polyene natural products via iterative cross-coupling. *J. Am. Chem. Soc.* **130,** 466–468 (2008).
47. Wang, C. & Glorius, F. Controlled iterative cross-coupling: on the way to the automation of organic synthesis. *Angew. Chem. Int. Ed.* **48,** 5240–5244 (2009).
48. Woerly, E. M., Roy, J. & Burke, M. D. Synthesis of most polyene natural product motifs using just 12 building blocks and one coupling reaction. *Nat. Chem.* **6,** 484–491 (2014).
49. Negishi, E., Tan, Z., Liang, B. & Novak, T. An efficient and general route to reduced polypropionates via Zr-catalyzed asymmetric C—C bond formation. *Proc. Natl Acad. Sci. USA* **101,** 5782–5787 (2004).
50. Albert, B. J. & Yamamoto, H. A triple-aldol cascade reaction for the rapid assembly of polyketides. *Angew. Chem. Int. Ed.* **49,** 2747–2749 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank EPSRC (EP/I038071/1) and the European Research Council (FP7/2007-2013, ERC grant no. 246785) for financial support. M.B. thanks the EPSRC-funded Bristol Chemical Synthesis Centre for Doctoral Training (EP/G036764/1) and Novartis for a PhD studentship. We wish to thank C. Woodall for assistance with X-ray analysis and E. Bozoki for assistance with preparative high-performance liquid chromatography purification.

Author Contributions V.K.A. designed the project. M.B. conducted and designed the experiments and analysed the data. S.E. performed computational studies and analysed the data with J.N.H. J.R.B. and S.P.B. performed the NMR experiments and analysed the data with C.P.B. M.P.W. conducted the preliminary experiments with lithiated carbamates. S.B. optimized the recrystallization conditions for stannane **5**. J.W.D. supervised M.B. while working at Novartis. V.K.A., M.B., J.N.H. and C.P.B. wrote the manuscript.

Author Information X-ray crystallographic data have been deposited in the Cambridge Crystallographic Data Centre database with accession numbers CCDC 993442 (**12**), CCDC 993443 (**15**) and CCDC 993441(**17**). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to V.K.A. (v.aggarwal@bristol.ac.uk), C.P.B. (craig.butts@bristol.ac.uk) for NMR work or J.N.H. (jeremy.harvey@bristol.ac.uk) for computational work.

doi:10.1038/nature13724

A synaptic and circuit basis for corollary discharge in the auditory cortex

David M. Schneider[1]*, Anders Nelson[1]* & Richard Mooney[1]

Sensory regions of the brain integrate environmental cues with copies of motor-related signals important for imminent and ongoing movements. In mammals, signals propagating from the motor cortex to the auditory cortex are thought to have a critical role in normal hearing and behaviour, yet the synaptic and circuit mechanisms by which these motor-related signals influence auditory cortical activity remain poorly understood. Using *in vivo* intracellular recordings in behaving mice, we find that excitatory neurons in the auditory cortex are suppressed before and during movement, owing in part to increased activity of local parvalbumin-positive interneurons. Electrophysiology and optogenetic gain- and loss-of-function experiments reveal that motor-related changes in auditory cortical dynamics are driven by a subset of neurons in the secondary motor cortex that innervate the auditory cortex and are active during movement. These findings provide a synaptic and circuit basis for the motor-related corollary discharge hypothesized to facilitate hearing and auditory-guided behaviours.

In a wide variety of sensory systems, including the auditory system, copies of motor signals (that is, corollary discharge signals) are used to modulate sensory processing in a movement-dependent manner[1–7]. In humans, evidence of this motor influence includes modulation of auditory cortical activity during vocalization and music-related manual gestures[8–10]. More broadly, corollary discharge signals are theorized to facilitate hearing during acoustic behaviours, to convey predictive signals important for complex forms of motor learning, such as speech and music, and to drive auditory hallucinations in certain pathological states[11–15].

While motor-related signals are likely to influence auditory processing at many sites in the brain[1,11,16,17], those operating at cortical levels are apt to be especially important to learning acoustic behaviours and generating auditory hallucinations[1,11,12,17,18]. Although the synaptic and circuit origins of corollary discharge signals in the auditory cortex remain enigmatic, a direct projection from the motor cortex to the auditory cortex is a common feature of the mammalian brain[19–21], providing a substrate for conveying corollary discharge signals to the auditory cortex. Moreover, heightened motor cortical activity correlates with auditory cortical suppression in humans[22], and activating motor cortical synapses in the auditory cortex suppresses tone-evoked auditory cortical responses in the anaesthetized mouse[20].

Despite the widespread observation that movement can modulate auditory cortical activity and the presumed role of the motor cortex in driving this modulation, the synaptic and circuit mechanisms by which the motor cortex influences auditory cortical activity during movement remain unresolved. Identifying these mechanisms requires integrating high-resolution electrophysiology techniques with circuit dissection strategies in freely behaving animals to establish causal links among synapses, circuits and behaviour. Here we combine *in vivo* intracellular physiology with optogenetic circuit manipulations in freely behaving mice to identify a synaptic signature of movement in the auditory cortex and to elucidate local and long-range circuits that modulate auditory cortical activity during movement.

Movement modulates auditory cortex

To begin to study how movement modulates auditory cortical processing at synaptic and circuit levels, we used a miniature motorized microdrive to make sharp electrode intracellular current-clamp recordings from putative auditory cortical excitatory neurons of freely behaving mice (Fig. 1a, b; Extended Data Fig. 1a; cells were classified as putative excitatory cells based on their intrinsic properties, spontaneous and DC-evoked action potential patterns and, in a subset of neurons, intracellular staining and post-hoc visualization). This approach permitted intermediate to long duration recordings (mean recording duration: 14.25 min, up to 155 min) accompanied by simultaneous video monitoring of head and body movements (Extended Data Fig. 1b–d). Immediately before and during a variety of movements including locomotion, head movements, and other body movements such as grooming, auditory excitatory neurons exhibited decreased variability in their sub-threshold membrane potential fluctuations and a slight depolarization (Fig. 1c, e). We also were able to elicit vocalizations in a subset of experiments (5 neurons in 3 mice) and although vocalizations were always accompanied by other head and body movements, the membrane potential dynamics during vocalization were indistinguishable from those observed during head movements, body movements, and translocation (Extended Data Fig. 2a, b). Therefore, the sub-threshold dynamics of mouse auditory cortical excitatory cells change in a stereotyped manner before and during the execution of a wide variety of natural behaviours.

To more precisely interrogate the relative timing between movement initiation and movement-related signals in the auditory cortex and to facilitate stimulus presentation and optogenetic manipulation of neuronal activity, we developed a head-fixed preparation for recording intracellular auditory cortical activity from mice free to move or rest on a quiet, non-motorized treadmill (Fig. 1b, Extended Data Fig. 1e–i, Extended Data Fig. 2c–e). The changes in membrane potential dynamics in the auditory cortex of head-fixed mice during treadmill locomotion, grooming, facial movements, posturing and forelimb movements were indistinguishable from those we observed in unrestrained, freely behaving mice, were consistent across superficial and deeper cortical layers, and also resembled state-dependent changes that have been observed in the mouse auditory cortex[23] (Fig. 1d, e, Extended Data Fig. 2f–h). By monitoring the onset and duration of locomotor bouts, we found that changes in membrane potential dynamics of auditory cortical excitatory neurons preceded locomotion on average by >200 ms and typically outlasted locomotion by a similar

[1]Department of Neurobiology, Duke University School of Medicine, Durham, North Carolina 27710, USA.
*These authors contributed equally to this work.

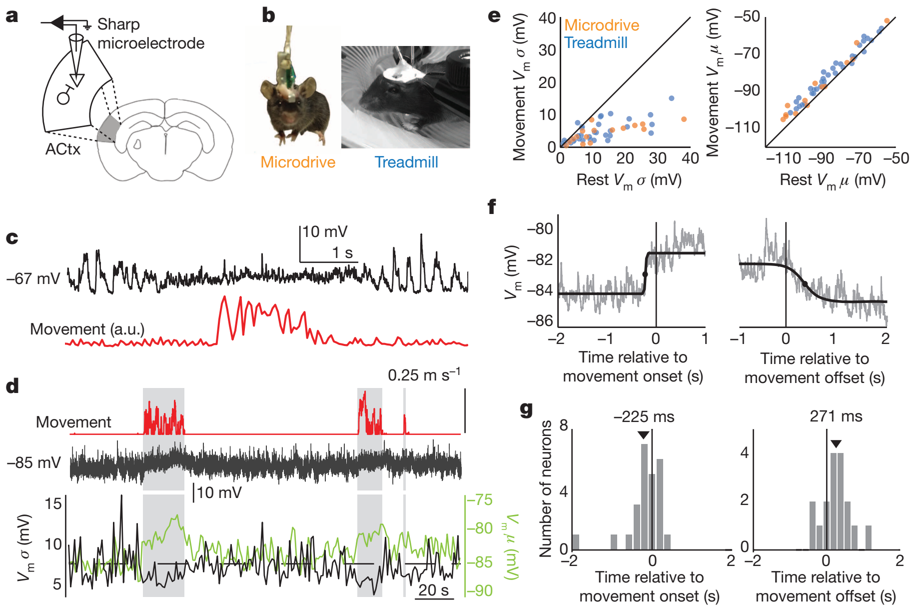


Figure 1 | Movement modulates membrane potential dynamics of auditory cortical neurons. a, Schematic showing sharp microelectrode current-clamp recording from an auditory cortical (ACtx) excitatory neuron in the behaving mouse. **b**, Video stills showing mouse with intracellular microdrive (left) and head-fixed mouse on treadmill (right). **c**, Membrane potential (top) of an auditory cortical neuron in an unrestrained mouse during rest and a brief movement (bottom; a.u., arbitrary units). **d**, Membrane potential (middle) of an auditory cortical neuron in a head-restrained mouse during rest and long bouts of locomotion on a treadmill (top). The bottom panel depicts the variance (σ) and mean (μ) of membrane potential (V_m) across time. **e**, Membrane potential variance and mean during rest versus movement ($n = 16/37$ cells for microdrive/treadmill, $P < 0.001$ for variance and mean, paired t-test). **f**, Membrane potential of an example neuron relative to movement onset and offset, averaged across 20 movements. Black lines show sigmoidal fit and the black dots show half rise/fall times. **g**, Histogram of the lag between membrane depolarization and movement onset (left, $P = 0.06$, t-test) and membrane hyperpolarization and movement offset (right, $P = 0.01$, t-test). Black arrows indicate population means ($n = 25$ cells). Statistical details in Methods.

duration (Fig. 1f, g). The finding that changes in auditory cortical membrane potential dynamics preceded locomotion onset by several to many hundreds of milliseconds indicates they cannot be caused by sensory reafference generated by the ensuing movements, and instead could reflect a motor-related signal. Moreover, movement-related changes in auditory cortical membrane potential dynamics persisted in the presence of broadband noise that was sufficiently loud to mask tone-evoked responses, further supporting the motor-related nature of these signals (Extended Data Fig. 3).

Suppression is local to cortex

In the visual cortex of the mouse, pyramidal neurons also display less variable and more positive membrane potentials before and during locomotion[5], and these changes are accompanied by a heightened responsiveness to visual stimuli[5,24,25]. In contrast, in the auditory cortex sound-evoked action potential responses are often suppressed during movement and during task engagement[11,23,26]. To determine whether stimulus-evoked responses in the auditory cortex of the mouse were enhanced or suppressed during movement, we presented tones at a neuron's best frequency during periods of rest and movement. In contrast to findings in the visual cortex, stimulus (that is, tone)-evoked synaptic responses of auditory cortical excitatory neurons were significantly diminished during movements (Fig. 2a, b, Extended Data Fig. 4a). Furthermore, in a small subset of neurons ($n = 4$) for which we measured tone-evoked responses during rest and movement at several different stimulus frequencies, we observed suppression during movement at all frequencies tested (Extended Data Fig. 4b).

Motor-related signals could modulate sound-evoked responses at multiple sites, spanning from the tympanic membrane to the auditory cortex[11,27], leading us to quantify the degree to which the motor-related suppression of tone-evoked responses occurs locally within the auditory cortex. We used viral vectors to express channelrhodopsin-2 (ChR2) in auditory thalamic neurons and subsequently placed an optical fibre over the auditory cortex to enable selective optical activation of thalamocortical axons (Fig. 2c, Extended Data Fig. 5a). Whereas acoustic stimulation activates the auditory system from the periphery to the cortex, this optogenetic approach allowed us to effectively bypass the ascending auditory pathway before the thalamocortical projection and to thus isolate the component of motor-related suppression that occurs within the auditory cortex. Movement was accompanied by a strong suppression of optogenetically evoked synaptic potentials recorded in auditory cortical excitatory neurons, consistent with a cortical locus (Fig. 2d, e). We also found that the degree of suppression of optogenetically evoked activity during movement was linearly related to the degree of suppression of tone-evoked activity during movement (slope = 0.6), indicating that approximately 60 per cent of the movement-related suppression of tone-evoked responses in auditory cortical excitatory neurons arises through mechanisms local to the auditory cortex (Fig. 2f).

Together, these findings indicate that motor-related signals act in the auditory cortex to suppress sensory responses in auditory cortical excitatory neurons, pointing to the engagement of local inhibition. The intracellular methods used here allowed us to measure several properties of excitatory neurons, including their intrinsic excitability, input resistance, and the reversal potential of their motor-related synaptic currents, that can be used to further determine whether this inhibition acts at a pre- or postsynaptic locus. Injecting positive and negative current pulses through the recording electrode revealed that movement was accompanied by decreased excitability and input resistance of auditory cortical excitatory neurons, and these changes also could be detected before movement onset (Fig. 2g–j, Extended Data Fig. 5b–d). Additionally, tonic current injection was used to vary the resting potential of a subset of neurons. The movement-related change in mean membrane potential reversed in sign at approximately −72 mV, which was 3 mV depolarized relative to the average resting V_m and close to the chloride equilibrium potential reported for mouse auditory cortical pyramidal neurons[28] (Extended Data Fig. 6).

All of these features indicate that motor-related signals suppress auditory cortical excitatory cells via postsynaptic inhibition, but do not exclude the possibility that presynaptic inhibition is also involved. Specifically,

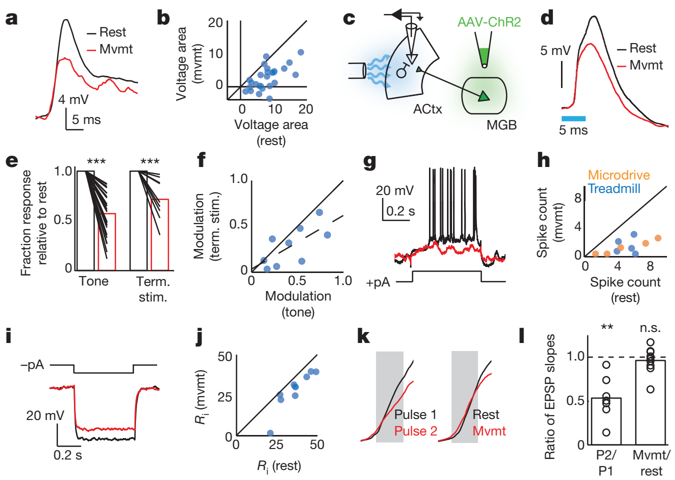


Figure 2 | Auditory cortical excitatory neurons are suppressed during movement. **a**, An example neuron's response to a preferred tone during rest (black) and movement (mvmt; red). **b**, The voltage area response of multiple neurons to preferred tone stimulus during rest versus movement ($n = 27$, $P < 0.001$, paired t-test). **c**, Schematic showing viral infection of AAV-ChR2 into medial geniculate body (MGB) and optogenetic activation of ChR2$^+$ axon terminals in the auditory cortex. **d**, Responses of an example neuron to optogenetic stimulation of thalamocortical terminals during rest (black) and movement (red). Blue bar indicates duration of light stimulation. **e**, Normalized responses to preferred tone stimulus (left, $n = 27$, t-test) and thalamocortical terminal stimulation (term. stim.; right, $n = 9$, t-test) during rest (black bars) and movement (red bars). **f**, Modulation of tone-evoked versus thalamocortical terminal stimulation. Modulation was defined as $(1 - R_{mvmt}/R_{rest})$, where R_{mvmt} and R_{rest} were the peak response during movement and rest, respectively. Dashed line shows the linear regression ($n = 9$, $r = 0.69$). **g**, Evoked response of an example neuron to intracellular positive current injection during rest (black) and movement (red). **h**, Number of spikes evoked by positive current injection during rest versus movement ($n = 5/5$ for microdrive/treadmill, $P < 0.001$, paired t-test). **i**, Evoked voltage response of an example neuron to intracellular negative current injection during rest (black) and movement (red). **j**, Input resistance (R_i) during rest versus movement ($n = 10$, $P < 0.001$, paired t-test). **k**, Initial phase of the EPSP evoked by optogenetic stimulation of thalamic terminals in the auditory cortex. Left panel shows the responses to two sequential pulses delivered 50 ms apart. Right panel shows the responses to pulses delivered during rest and movement. **l**, Ratio of the EPSP slopes measured during paired-pulse stimulation (pulse 2 vs pulse 1, P2/P1; $n = 7$, t-test) and during movement vs rest ($n = 9$, t-test). n.s., not significant; $^{**}P < 0.01$, $^{***}P < 0.001$. Statistical details in Methods.

task-engagement has been shown to augment auditory thalamic neuron activity[29,30], which could depress thalamic terminals in the auditory cortex. To explore this possibility we first established that optogenetically stimulating auditory thalamocortical synapses at >20 Hz was sufficient to decrease the onset slope of the evoked excitatory postsynaptic potential (EPSP), which provides a postsynaptic readout of presynaptic depression (Fig. 2k, l)[31,32]. We then measured the onset slope of optogenetically evoked thalamocortical synaptic potentials, and found no difference in the rising slope of the EPSP across movement and rest conditions, indicating that thalamic terminals in the auditory cortex are not depressed during movement (Fig. 2k, l). Therefore, motor-related suppression of tone-evoked responses in the auditory cortex involves postsynaptic inhibition on excitatory neurons.

PV$^+$ neurons are active during movement

To explicitly test if auditory cortical inhibitory neurons were recruited by a motor-related signal and involved in movement-related changes in excitatory neuron activity, we monitored the spiking activity of a large population of neurons using a multi-electrode array inserted across a broad expanse of the auditory cortex in mice engineered to express ChR2

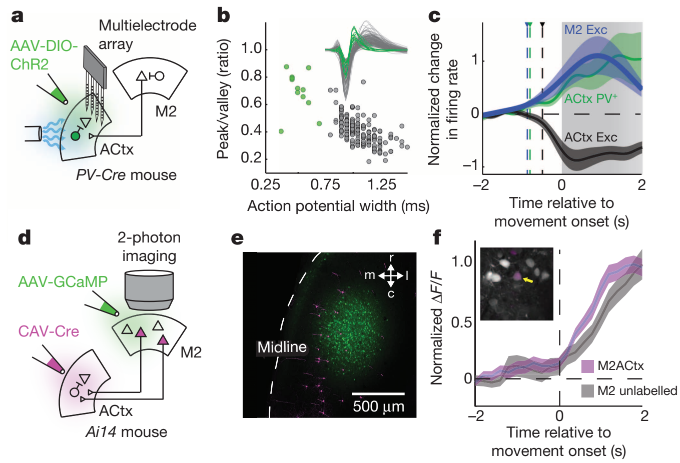


Figure 3 | Auditory cortical PV$^+$ interneurons and M2$_{ACtx}$ neurons are active during movement. **a**, Schematic showing viral infection of PV-Cre mice with a Cre-dependent ChR2 construct. **b**, Width and peak-to-valley ratio of action potentials of excitatory (black, $n = 173$) and PV$^+$ (green, $n = 12$) auditory cortical neurons. Inset shows average action potential of every neuron. **c**, Average spiking activity of excitatory (black) and PV$^+$ (green) populations in the auditory cortex and excitatory M2 neurons (blue, $n = 90$) relative to movement onset, normalized to spontaneous firing levels during rest. Triangles and dashed vertical lines show time of significant deviation from resting. All movements lasted at least 1 s, and 80 per cent of movements persisted for at least 2 s, as indicated by gradation of grey bar. **d**, Schematic showing Cre-dependent expression of tdTomato in M2$_{ACtx}$ neurons, injection of M2 with GCaMP6s, and two-photon calcium imaging. **e**, tdTomato$^+$ and tdTomato$^-$ M2 neurons expressing GCaMP6s in M2 *ex vivo*. **f**, Change in fluorescence of tdTomato$^+$ ($n = 7$) and tdTomato$^-$ ($n = 23$) M2 neurons aligned to movement onset. Inset shows a representative imaging region in M2 with tdTomato$^+$ and tdTomato$^-$ M2 neurons expressing GCaMP6s. Data show mean ± s.e. Statistical details in Methods.

in parvalbumin-positive (PV$^+$) or other GABAergic neurons (Fig. 3a, Extended Data Fig. 7a, b, PV–ChR2 and vesicular GABA transporter (VGAT)–ChR2 mice, respectively)[33]. We then classified neurons as PV$^+$ cells, inhibitory cells or excitatory cells on the basis of their action potential shapes (Fig. 3b, Extended Data Fig. 7c) and whether they were excited or suppressed by blue light. Before and during movements, the firing rates of PV$^+$ cells and other fast-spiking interneurons increased, whereas the firing rates of putative excitatory neurons decreased (Fig. 3c, Extended Data Fig. 7d, e). As a population, inhibitory neuron firing rates increased well before movement onset and also before the firing rates of auditory cortical excitatory cells decreased below their baseline levels (−805 ms for PV$^+$ cells, −605 ms for VGAT$^+$ cells, and −490 ms for excitatory cells, relative to movement onset). These experiments indicate that in the auditory cortex, motor-related signals excite PV$^+$ interneurons, which in turn postsynaptically inhibit excitatory cells to suppress their spontaneous activity and stimulus-evoked responses.

M2$_{ACtx}$ neurons are active during movement

What is the source of motor-related signals in the auditory cortex? Anatomical tracing studies in the mouse show that the auditory cortex receives input from several motor-related regions, including the cingulate cortex, primary motor cortex, and secondary motor cortex (M2), the last of which when optogenetically activated can drive strong feedforward inhibition in the auditory cortex mediated in part by PV$^+$ interneurons[20] (Extended Data Fig. 8a). Moreover, a subset of M2 neurons have branching axons that innervate the auditory cortex and the brainstem, providing an anatomical substrate for conveying motor-related signals to the auditory cortex[20]. If M2 is a source of motor-related signals in freely behaving mice, M2 excitatory cells should increase their firing rate before movement-related changes in auditory cortical activity. Using an extracellular array that spanned a large region of M2 in mice freely behaving

on a treadmill, we determined that M2 excitatory neuron action potential activity increased before a variety of movements, including locomotion (Fig. 3c; −870 ms, relative to movement onset). Notably, pre-movement increases in M2 activity also preceded changes in auditory cortical spiking activity and membrane potential dynamics and, at movement offset, M2 activity declined to baseline levels with a time course similar to movement offset-related changes in the auditory cortex (Extended Data Fig. 8b, c; also see Fig. 1g).

A remaining issue is whether M2 neurons that extend axons to the auditory cortex (that is, $M2_{ACtx}$ cells) display movement-related dynamics similar to the general population of M2 cells recorded here using a multi-electrode array. To resolve this issue, we used an intersectional strategy to selectively express a red fluorescent reporter in $M2_{ACtx}$ cells and then used viral methods to express the genetically encoded calcium indicator GCaMP6s[34] in broad fields of M2 neurons (Fig. 3d, e). Two-photon calcium imaging of GCaMP6s-expressing M2 neurons in head-fixed mice running on a treadmill revealed that $M2_{ACtx}$ cells exhibited movement-related increases in fluorescence with a time course indistinguishable from that of the general M2 population (Fig. 3f). Thus, $M2_{ACtx}$ cells are a source of motor-related signals that could be transmitted to the auditory cortex.

$M2_{ACtx}$ terminals drive movement-like dynamics

To begin to test whether $M2_{ACtx}$ cells can account for changes in auditory cortical dynamics like those observed during movement, we assessed whether activating M2 terminals in the auditory cortex of resting mice was sufficient to induce movement-like membrane potential dynamics in auditory cortical excitatory neurons. Following viral infection of AAV-ChR2 in M2 (Extended Data Fig. 8d), optically activating $ChR2^+$ M2 terminals in the auditory cortex of resting mice decreased the membrane potential variability and tone-evoked responses of excitatory cells, and also resulted in a slight depolarization, highly similar to the effects of movement (Fig. 4a, b, e–g). One potential concern is that optogenetic activation of M2 terminals in the auditory cortex triggers antidromic propagation of action potentials and thus excites other targets of $M2_{ACtx}$ cells, some of which may also innervate the auditory cortex. Two observations argue against such an indirect mechanism. First, optogenetic activation of M2 terminals was equally efficacious in modulating auditory cortical dynamics when M2 cell bodies were pharmacologically silenced with the sodium channel blocker tetrodotoxin (TTX; Fig. 4c–g). Second, the onset of changes in auditory cortical dynamics following optogenetic activation of M2 terminals occurred more rapidly (<7 ms) than the antidromic propagation time from auditory cortex to M2 (~12 ms) (Fig. 4h). Therefore, activating M2 terminals within the auditory cortex is sufficient to induce movement-like auditory cortical dynamics without concomitant recruitment of indirect pathways.

M2 is necessary for motor-related dynamics

These experiments raise the possibility that ongoing M2 activity is necessary for maintaining movement-related synaptic dynamics in the auditory cortex. To test this idea, we unilaterally and selectively silenced M2 excitatory neuron cell bodies during locomotion by optogenetically activating M2 inhibitory neurons in VGAT–ChR2 mice (Fig. 5a–c, Extended Data Fig. 8e). Silencing either ipsilateral or contralateral M2 (relative to the auditory cortical recording site) was sufficient to arrest movements ~500 ms after light onset (Fig. 5d–f). Notably, silencing ipsilateral M2 rapidly (~70 ms after laser onset) restored rest-like membrane potential dynamics in the auditory cortex (Fig. 5f–h), and this reversion to a rest-like state always preceded movement offset (Fig. 5f, Extended Data Fig. 8f). In contrast, silencing contralateral M2 did not lead to changes in auditory cortical excitatory neurons until after movement offset, effectively recapitulating the time course observed after spontaneous movement cessation (Fig. 5f, also see Fig. 1g, h). These differential effects of silencing ipsilateral versus contralateral M2 on auditory cortical dynamics are consistent with a previous anatomical finding that the projection from M2 to auditory cortex is almost completely ipsilateral[20].

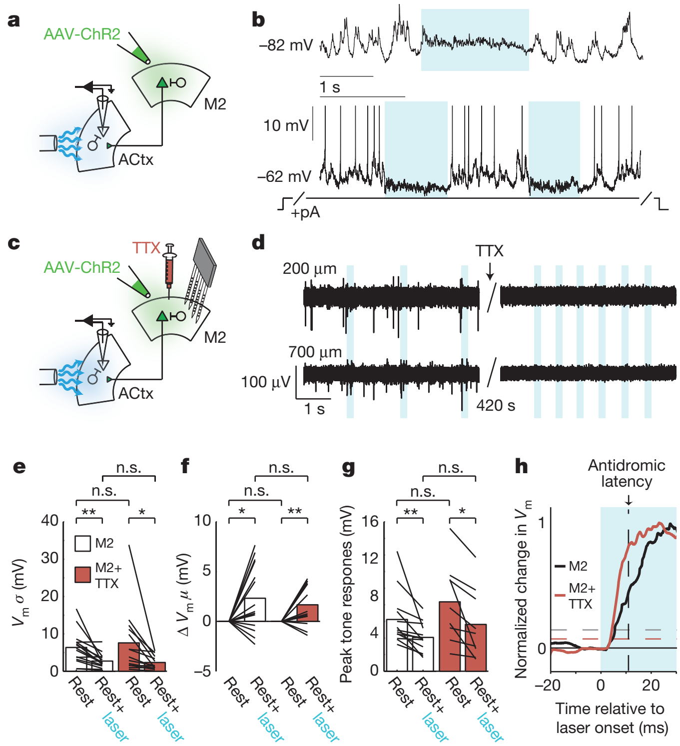


Figure 4 | M2 axon terminals in the auditory cortex are sufficient to produce movement-like auditory cortical dynamics during rest. **a**, Schematic showing intracellular recording in the auditory cortex during optogenetic activation of $ChR2^+$ M2 terminals. **b**, Optogenetic stimulation of M2 terminals in the auditory cortex causes a slight depolarization and decreased variability (top), and during tonic depolarization, M2 terminal stimulation suppresses spontaneous spiking and hyperpolarizes neurons (bottom, spikes truncated). **c**, Schematic showing intracellular recording in auditory cortex during optogenetic activation of $ChR2^+$ M2 terminals and multi-electrode array recordings in M2 during pharmacological silencing of M2 cell bodies with the sodium channel blocker, TTX. **d**, Left panels show superficial and deep recordings in M2 with spontaneous spikes (top and bottom) and antidromic spikes evoked by optogenetic stimulation of $M2_{ACtx}$ terminals (blue). Right panels show the abolition of spontaneous and antidromic spiking in M2 after TTX application. **e–g**, M2 terminal stimulation leads to decreases in membrane potential variability (**e**, $n = 15/14$, paired t-test), a slight depolarization (**f**, $n = 15/14$, t-test), and decreased tone-evoked responses (**g**, $n = 13/9$, paired t-test), with and without M2 cell bodies inactivated with TTX (n = number of cells recorded without/with TTX; legend in **e** applies to **e–g**). **h**, Normalized average change in membrane potential after M2 terminal stimulation with (red, $n = 15$) and without (black, $n = 14$) M2 cell bodies inactivated. Vertical black dashed line shows the latency of an antidromic spike travelling from the auditory cortex to M2. Horizontal dashed lines indicate significant depolarizations relative to baseline. $*P < 0.05$, $**P < 0.01$. Statistical details in Methods.

The observation that silencing the ipsilateral M2 could restore rest-like auditory cortical dynamics several hundred milliseconds before movement offset allowed us to determine whether unilaterally silencing M2 was sufficient to enhance tone-evoked responses in the ipsilateral auditory cortex while the mouse was still moving. Presenting tones during the initial phase of M2 suppression, when auditory cortical membrane potential dynamics had transitioned to a rest-like state but before locomotion offset, revealed a strong (~40 per cent) recovery of tone-evoked responses (Fig. 5i, j). We also noted that silencing M2 excitatory cell bodies in the resting mouse could slightly enhance tone-evoked responses in the auditory cortex, consistent with the idea that spontaneous activity in M2 of resting, awake mice exerts a weak suppressive effect on auditory cortical responsiveness (Fig. 5i, j). Together, these experiments dissociate the motor-related modulations we observe in the auditory cortex from movement, and show that activity in ipsilateral M2 plays a critical role

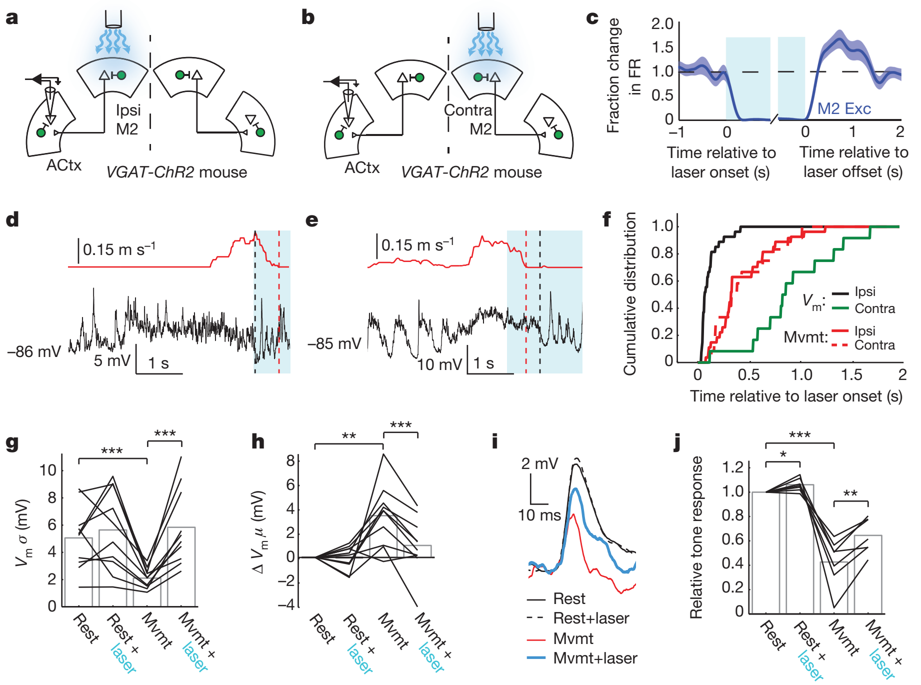


Figure 5 | M2 activity is necessary to sustain movement-related dynamics in the auditory cortex. **a**, **b**, Schematics showing intracellular recording in auditory cortex while silencing ipsilateral (**a**) or contralateral (**b**) M2. **c**, Average spiking activity (mean ± s.e.) of a population of M2 excitatory neurons ($n = 66$) before, during and after optogenetic activation of M2 inhibitory neurons (FR, firing rate). **d**, Membrane potential dynamics of example auditory cortical excitatory neuron during rest, during movement (red trace) and during movement with optogenetic suppression of ipsilateral M2 excitatory neurons (blue bar). **e**, As in **d**, but while silencing contralateral M2. **f**, Transition to rest-like membrane potential dynamics precedes movement offset with ipsilateral M2 silencing ($n = 27$, $P < 0.001$, two-sample Kolmogorov–Smirnov (KS) test) but follows movement offset with contralateral M2 silencing ($n = 12$, $P < 0.05$, two-sample KS test). **g**, **h**, Membrane potential variance (**g**, $n = 10$, paired t-test) and mean (**h**, $n = 10$, paired t-test) of auditory cortical excitatory neurons during rest and movement with and without optogenetic suppression of ipsilateral M2. **i**, Tone-evoked responses of an example neuron during rest and movement, with and without optogenetic suppression of ipsilateral M2. **j**, Tone-evoked responses of auditory cortical excitatory neurons during rest and movement, with and without optogenetic suppression of ipsilateral M2 ($n = 7$, paired t-test). $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. Statistical details in Methods.

in driving movement-like synaptic dynamics and controlling the gain of sensory-evoked responses in the auditory cortex.

Discussion

Projections from motor cortex to the auditory cortex are an architectural feature common to many mammalian species[19–21,35,36], including humans and other primates, and are thought to convey information critical for learning and executing complex behaviours, including speech and musicianship. Although movement-related modulation of auditory cortical activity has been detected in monkey and human auditory cortex during a variety of behaviours[1,10,11,18,37], a direct role for the motor cortex in these modulatory processes was untested. By applying a wide range of electrophysiological, optical, optogenetic and pharmacological methods in the freely behaving mouse, this study identifies a postsynaptic inhibitory signature of motor action within auditory cortex, a local source of this inhibition, and a long-range motor-to-auditory cortical circuit that engages this local inhibitory mechanism to suppress tone-evoked responses during movement. We found that a wide variety of natural movements strongly suppresses the spontaneous and tone-evoked synaptic activity of auditory cortical excitatory cells and that a substantial fraction of this suppression is mediated through a postsynaptic mechanism involving increased local inhibition via PV^+ interneurons. This mechanism contrasts with a disinhibitory mechanism implicated in locomotion-dependent increases in visual cortical responses[38], with a parallel negative rescaling of excitatory and inhibitory synaptic drive that has been advanced to account for state-dependent changes in auditory cortical responsiveness[23], and with presynaptic depression driven by state-dependent increases in thalamic activity[30]. Moreover, our observation that this suppression precedes movement onset and persists in masking noise strongly implicates a motor-related signal, rather than sensory reafference or attentional mechanisms[26]. Finally, the finding that movement can suppress ChR2-evoked auditory thalamocortical responses indicates that motor-related suppression of tone-evoked responses is not simply a consequence of peripheral masking by movement-related noise[27,39].

The present findings establish that direct ipsilateral projections from M2 to the auditory cortex are sufficient to account for movement-related auditory cortical dynamics and that activity in the ipsilateral M2 is also necessary to sustain these dynamics during movement. However, M2 and the auditory cortex are embedded in complex networks, a consequence of which is that, in addition to directly influencing auditory cortical processing, M2 could also act indirectly through or in concert with neuromodulatory cell groups[40] to modulate auditory cortical dynamics. These findings add to a growing body of evidence that motor-related signals, including those arising from motor cortical regions, can strongly modulate the stimulus-evoked responsiveness of sensory cortical neurons[1,2,4,5,24,25]. Notably, whereas the gain of stimulus-evoked responses in visual cortical pyramidal cells increases with locomotion[5,25], perhaps to compensate for increased visual flow, auditory cortical responses to tones decreased during movement. This suppressive effect, which resembles corollary discharge signals described in the auditory systems of animals ranging from insects to humans[6,41], may reflect a general strategy where motor-related signals transiently dampen sensitivity to predictable low-intensity sounds, enabling auditory neurons to maintain responsiveness to unexpected high-intensity stimuli[23,42]. Finally, motor–auditory cortical circuitry is implicated in various forms of abnormal hearing, including tinnitus[43] and auditory hallucinations[12,44], motivating future studies to investigate structural and functional changes in this circuitry in appropriate animal models.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 18 March; accepted 31 July 2014.

Published online 27 August 2014.

1. Eliades, S. J. & Wang, X. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J. Neurophysiol.* **89,** 2194–2207 (2003).
2. Lee, S., Kruglikov, I., Huang, Z. J., Fishell, G. & Rudy, B. A disinhibitory circuit mediates motor integration in the somatosensory cortex. *Nature Neurosci.* **16,** 1662–1670 (2013).
3. Maimon, G., Straw, A. D. & Dickinson, M. H. Active flight increases the gain of visual motion processing in *Drosophila. Nature Neurosci.* **13,** 393–399 (2010).
4. Petreanu, L. *et al.* Activity in motor-sensory projections reveals distributed coding in somatosensation. *Nature* **489,** 299–303 (2012).
5. Polack, P. O., Friedman, J. & Golshani, P. Cellular mechanisms of brain state-dependent gain modulation in visual cortex. *Nature Neurosci.* **16,** 1331–1339 (2013).
6. Poulet, J. F. & Hedwig, B. The cellular basis of a corollary discharge. *Science* **311,** 518–522 (2006).
7. Sommer, M. A. & Wurtz, R. H. A pathway in primate brain for internal monitoring of movements. *Science* **296,** 1480–1482 (2002).
8. Flinker, A. *et al.* Single-trial speech suppression of auditory cortex activity in humans. *J. Neurosci.* **30,** 16643–16650 (2010).
9. Reznik, D., Henkin, Y., Schadel, N. & Mukamel, R. Lateralized enhancement of auditory cortex activity and increased sensitivity to self-generated sounds. *Nature Commun.* **5,** 4059 (2014).
10. Zatorre, R. J., Chen, J. L. & Penhune, V. B. When the brain plays music: auditory-motor interactions in music perception and production. *Nature Rev. Neurosci.* **8,** 547–558 (2007).
11. Eliades, S. J. & Wang, X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* **453,** 1102–1106 (2008).
12. Ford, J. M. & Mathalon, D. H. Corollary discharge dysfunction in schizophrenia: can it explain auditory hallucinations? *Int. J. Psychophysiol.* **58,** 179–189 (2005).
13. Hickok, G., Houde, J. & Rong, F. Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* **69,** 407–422 (2011).
14. Houde, J. F. & Jordan, M. I. Sensorimotor adaptation in speech production. *Science* **279,** 1213–1216 (1998).
15. Paus, T., Perry, D. W., Zatorre, R. J., Worsley, K. J. & Evans, A. C. Modulation of cerebral blood flow in the human auditory cortex during speech: role of motor-to-sensory discharges. *Eur. J. Neurosci.* **8,** 2236–2246 (1996).
16. Mukerji, S., Windsor, A. M. & Lee, D. J. Auditory brainstem circuits that mediate the middle ear muscle reflex. *Trends Amplif.* **14,** 170–191 (2010).
17. Schuller, G. Vocalization influences auditory processing in collicular neurons of the CF-FM-bat, *Rhinolophus ferrumequinum*. *J. Comp. Physiol.* **132,** 39–46 (1979).
18. Curio, G., Neuloh, G., Numminen, J., Jousmaki, V. & Hari, R. Speaking modifies voice-evoked activity in the human auditory cortex. *Hum. Brain Mapp.* **9,** 183–191 (2000).
19. Budinger, E. & Scheich, H. Anatomical connections suitable for the direct processing of neuronal information of different modalities via the rodent primary auditory cortex. *Hear. Res.* **258,** 16–27 (2009).
20. Nelson, A. *et al.* A circuit for motor cortical modulation of auditory cortical activity. *J. Neurosci.* **33,** 14342–14353 (2013).
21. Reep, R. L., Corwin, J. V., Hashimoto, A. & Watson, R. T. Efferent connections of the rostral portion of medial agranular cortex in rats. *Brain Res. Bull.* **19,** 203–221 (1987).
22. Ford, J. M. & Mathalon, D. H. Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. *J. Psychiatr. Res.* **38,** 37–46 (2004).
23. Zhou, M. *et al.* Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nature Neurosci.* **17,** 841–850 (2014).
24. Bennett, C., Arroyo, S. & Hestrin, S. Subthreshold mechanisms underlying state-dependent modulation of visual responses. *Neuron* **80,** 350–357 (2013).
25. Niell, C. M. & Stryker, M. P. Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* **65,** 472–479 (2010).
26. Otazu, G. H., Tai, L. H., Yang, Y. & Zador, A. M. Engaging in an auditory task suppresses responses in auditory cortex. *Nature Neurosci.* **12,** 646–654 (2009).
27. Carmel, P. W. & Starr, A. Acoustic and nonacoustic factors modifying middle-ear muscle activity in waking cats. *J. Neurophysiol.* **26,** 598–616 (1963).
28. Tan, A. Y. & Wehr, M. Balanced tone-evoked synaptic excitation and inhibition in mouse auditory cortex. *Neuroscience* **163,** 1302–1315 (2009).
29. Jaramillo, S., Borges, K. & Zador, A. M. Auditory thalamus and auditory cortex are equally modulated by context during flexible categorization of sounds. *J. Neurosci.* **34,** 5291–5301 (2014).
30. Poulet, J. F., Fernandez, L. M., Crochet, S. & Petersen, C. C. Thalamic control of cortical states. *Nature Neurosci.* **15,** 370–372 (2012).
31. Korn, H., Faber, D. S., Burnod, Y. & Triller, A. Regulation of efficacy at central synapses. *J. Neurosci.* **4,** 125–130 (1984).
32. Thomson, A. M., Deuchars, J. & West, D. C. Large, deep layer pyramid-pyramid single axon EPSPs in slices of rat motor cortex display paired pulse and frequency-dependent depression, mediated presynaptically and self-facilitation, mediated postsynaptically. *J. Neurophysiol.* **70,** 2354–2369 (1993).
33. Guo, Z. V. *et al.* Flow of cortical activity underlying a tactile decision in mice. *Neuron* **81,** 179–194 (2014).
34. Chen, T. W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499,** 295–300 (2013).
35. Alexander, G. E., Newman, J. D. & Symmes, D. Convergence of prefrontal and acoustic inputs upon neurons in the superior temporal gyrus of the awake squirrel monkey. *Brain Res.* **116,** 334–338 (1976).
36. Hackett, T. A., Stepniewska, I. & Kaas, J. H. Prefrontal connections of the parabelt auditory cortex in macaque monkeys. *Brain Res.* **817,** 45–58 (1999).
37. Phillips-Silver, J. & Trainor, L. J. Feeling the beat: movement influences infant rhythm perception. *Science* **308,** 1430 (2005).
38. Fu, Y. *et al.* A cortical circuit for gain control by behavioral state. *Cell* **156,** 1139–1152 (2014).
39. Horváth, J. & Burgyan, A. No evidence for peripheral mechanism attenuating auditory ERPs to self-induced tones. *Psychophysiology* **50,** 563–569 (2013).
40. Froemke, R. C., Merzenich, M. M. & Schreiner, C. E. A synaptic memory trace for cortical receptive field plasticity. *Nature* **450,** 425–429 (2007).
41. Crapse, T. B. & Sommer, M. A. Corollary discharge across the animal kingdom. *Nature Rev. Neurosci.* **9,** 587–600 (2008).
42. Buran, B. N., von Trapp, G. & Sanes, D. H. Behaviorally gated reduction of spontaneous discharge can improve detection thresholds in auditory cortex. *J. Neurosci.* **34,** 4076–4081 (2014).
43. Langguth, B. *et al.* Altered motor cortex excitability in tinnitus patients: a hint at crossmodal plasticity. *Neurosci. Lett.* **380,** 326–329 (2005).
44. Heinks-Maldonado, T. H. *et al.* Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Arch. Gen. Psychiatry* **64,** 286–296 (2007).

Acknowledgements We thank the members of the Mooney laboratory for discussions regarding experimental design and data analysis; S. Lisberger, F. Wang and S. Shea for their valuable comments on the manuscript; and M. Booze for technical support and animal husbandry. D.M.S. is a fellow of the Helen Hay Whitney Foundation; A.N. was supported by the Holland-Trice Graduate Fellowship in Brain Sciences; R.M. was supported by NIH grant NS079929.

Author Contributions D.M.S., A.N. and R.M. initiated the project and designed the experiments. D.M.S. performed electrophysiological, optogenetic, and pharmacological experiments in head-fixed mice. A.N. performed electrophysiological experiments in unrestrained mice, two-photon calcium imaging in head-fixed mice, immunohistochemistry, and imaging. D.M.S. and A.N. analysed the data. D.M.S., A.N. and R.M. prepared the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.M. (mooney@neuro.duke.edu).

ARTICLE

OPEN
doi:10.1038/nature13679

Gibbon genome and the fast karyotype evolution of small apes

A list of authors and their affiliations appears at the end of the paper

Gibbons are small arboreal apes that display an accelerated rate of evolutionary chromosomal rearrangement and occupy a key node in the primate phylogeny between Old World monkeys and great apes. Here we present the assembly and analysis of a northern white-cheeked gibbon (*Nomascus leucogenys*) genome. We describe the propensity for a gibbon-specific retrotransposon (LAVA) to insert into chromosome segregation genes and alter transcription by providing a premature termination site, suggesting a possible molecular mechanism for the genome plasticity of the gibbon lineage. We further show that the gibbon genera (*Nomascus*, *Hylobates*, *Hoolock* and *Symphalangus*) experienced a near-instantaneous radiation ~5 million years ago, coincident with major geographical changes in southeast Asia that caused cycles of habitat compression and expansion. Finally, we identify signatures of positive selection in genes important for forelimb development (*TBX5*) and connective tissues (*COL1A1*) that may have been involved in the adaptation of gibbons to their arboreal habitat.

Gibbons (Hylobatidae) are critically endangered[1] small apes that inhabit the tropical forests of southeast Asia (Fig. 1) and belong to the superfamily Hominoidea along with great apes and humans. In the primate phylogeny, gibbons diverged between Old World monkeys and great apes, providing a unique perspective from which to study the origins of hominoid characteristics.

Gibbons have several distinctive traits, the most striking of which is the unusually high number of large-scale chromosomal rearrangements in comparison to the inferred ancestral ape karyotype[2]. The four gibbon genera (*Nomascus*, *Hylobates*, *Hoolock* and *Symphalangus*) occupy different regions of southeast Asia and bear distinctive karyotypes, with diploid chromosome numbers ranging from 38 to 52 (Fig. 1). Given the relatively recent differentiation of these genera (4–6 million years ago (Myr ago)), this constitutes an extraordinarily fast rate of karyotype change.

In order to investigate the mechanisms behind the plasticity of the gibbon genome, understand the evolutionary relationships among the four extant gibbon genera and study the evolution of putatively functional sequences related to gibbon-specific adaptations, we sequenced and assembled the genome of a female northern white-cheeked gibbon (*Nomascus leucogenys*) named 'Asia'. The reference assembly (Nleu1.0) provides on average 5.7-fold Sanger read coverage over 2.9 gigabase pairs (Gb) (Table 1 and Supplementary Table ST1.1). Our quality assessment (Extended Data Fig. 1) confirmed its equivalence to other Sanger sequence-based non-human primate draft assemblies (such as the orangutan or rhesus macaque[3,4]) (Supplementary Information section S1, Supplementary Data Files 1 and 2). We also obtained ~15× whole-genome shotgun (WGS) short-read data (Illumina) for two individuals of each gibbon genus and high-coverage exome data (>60×) for two of the same individuals in order to derive error models for single nucleotide polymorphism (SNP) calls (Supplementary Information section S2; Supplementary Tables ST2.1–2.3).

Gibbon–human synteny breakpoints

Nleu1.0 scaffolds were aligned against the human reference (GRCh37) to be ordered and oriented into 26 chromosomes (Nleu3.0) under extensive guidance by cytogenetic data. The reshuffled nature of the gibbon genome was especially evident when human–gibbon chromosome alignments were compared with those between human and great apes, rhesus macaque (Old World monkey) and marmoset (New World monkey)

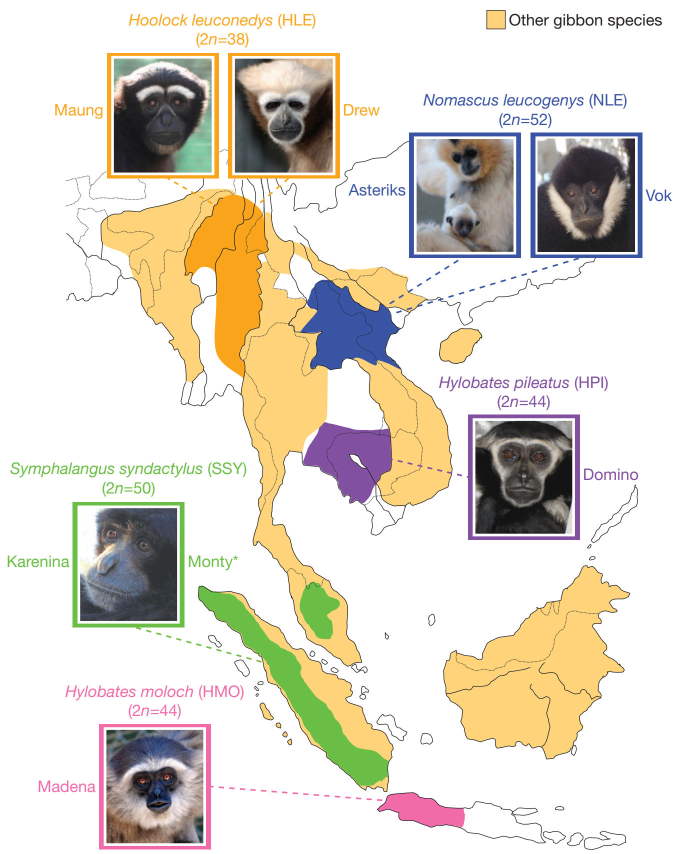


Figure 1 | Geographic distribution of gibbon species used in the study. We sequenced two individuals from each gibbon genus and two different species (*H. moloch* and *H. pileatus*) for the genus *Hylobates*. The extant geographic localization for each genus is illustrated on the map. Individuals in the photos are the ones sequenced in this study. The asterisk symbol indicates a deceased animal.

Table 1 | Gibbon assembly statistics

Assembly (Nleu1.0/nomLeu1)	
Total sequence length	2,936,052,603 bp
Ungapped length	2,756,591,777 bp
Total contig length	2.77 Gb (92.36%)
Number of contigs >1 kb	197,908
N50 contig length	35,148 bp
Number of scaffolds >3 kb	17,976
N50 scaffold length	22,692,035 bp
Average read depth	5.6×

(Fig. 2a). This higher rate of reshuffling applied only to large-scale chromosomal rearrangements (>10 megabases (Mb)), whereas smaller-scale rearrangements (10–100 kilobases (kb)) were comparable with other species (Fig. 2b) (Supplementary Information section S1).

We identified 96 gibbon–human synteny breakpoints in Nleu1.0 and classified them as to whether they could be defined at the base-pair level (class I, $n = 42$) or only narrowed to an interval due to greater complexity (class II, $n = 54$). As previously reported[5], breakpoints were significantly depleted of genes (Supplementary Fig. SF5.2 and Supplementary Data File 3) and breakpoint intervals contained a mixture of repetitive sequences that inserted exclusively into the gibbon genome[2,5,6] (Fig. 2c). To assess breakpoint segmental duplication content, we identified gibbon-specific segmental duplication using *in silico* methods followed by experimental validation (Extended Data Fig. 2, Supplementary Fig. SF3.1, Supplementary Information section S3 and Supplementary Data File 4). Of note, both gibbon-specific segmental duplication and gene family expansion analyses suggested the gibbon genome has not undergone a greater rate of duplication than other hominoids, further supporting a model in which accelerated evolution has been limited to gross chromosomal rearrangements (Supplementary Information section S6, Supplementary Fig. SF6.1).

Segmental duplication enrichment was the best predictor of gibbon–human synteny breakpoints, as shown through permutation analyses (P value < 0.0001); however, breakpoints were also enriched for *Alu* elements (Supplementary Table ST5.1; Supplementary Information section S5; Supplementary Fig. SF5.2). Although non-allelic homologous recombination between highly similar sequences can mediate large-scale rearrangements[7], the majority of gibbon chromosomal breakpoints bore signatures of non-homology based mechanisms (Fig. 2c). These included the insertion of non-templated sequences (2–51 nucleotides

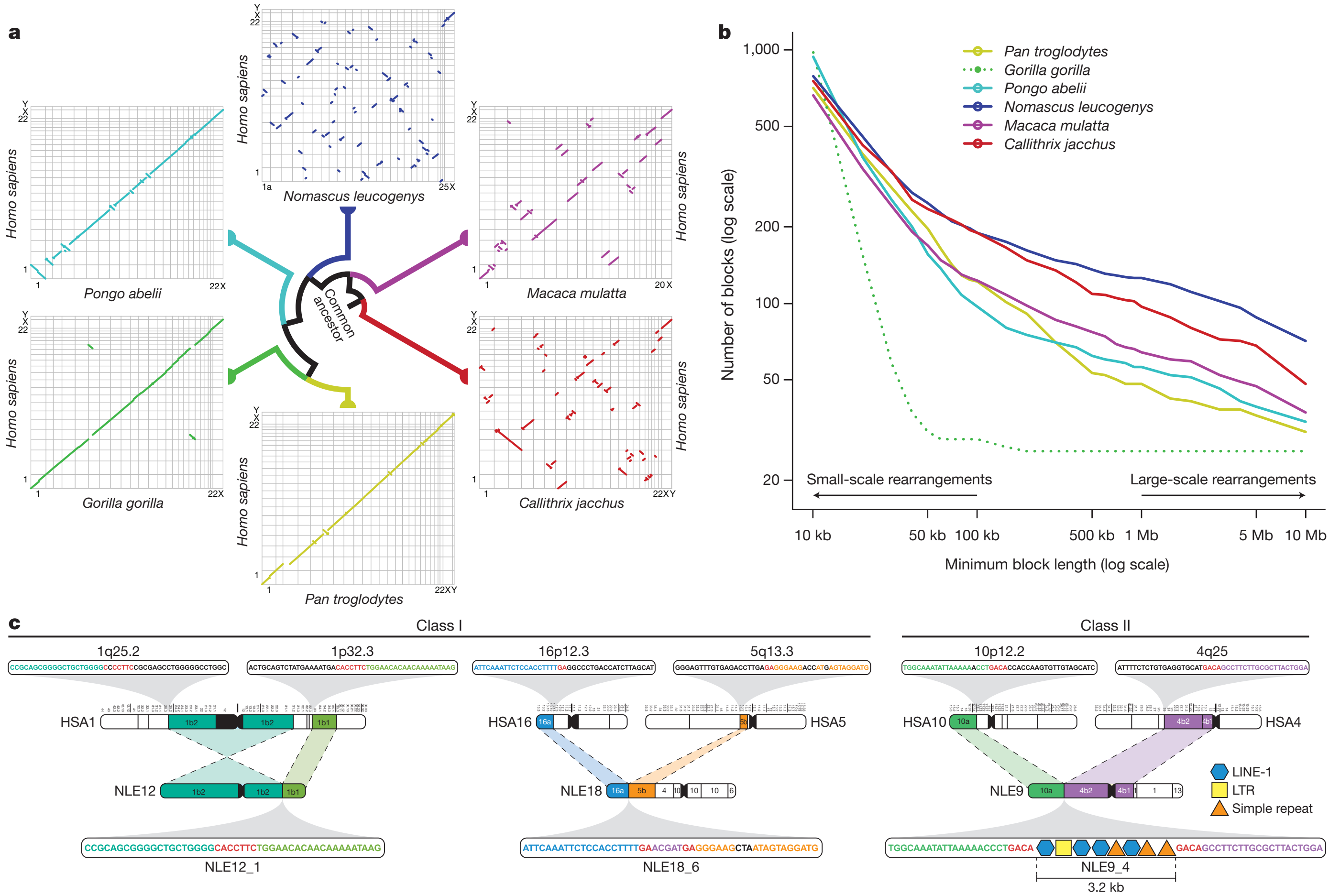


Figure 2 | Analysis of gibbon–human synteny and breakpoints. **a**, Oxford plots for human chromosomes (y axis) vs. chimpanzee, gorilla, orangutan, gibbon, rhesus macaque and marmoset chromosomes (x axis). Each line represents a collinear block larger than 10 Mb. The gibbon genome displays a significantly larger number of large-scale rearrangements than all the other species. In the gorilla plot, chromosomes 4 and 19 stand out as the product of a reciprocal translocation between chromosomes syntenic to human chromosomes 5 and 17. **b**, The graph shows the number of collinear blocks in primate genomes with respect to the human genome. The number of collinear blocks is a proxy for the number of rearrangements and decreases as the size of the blocks becomes larger. The gibbon genome has undergone a greater number of large-scale rearrangements; however, the number of small-scale rearrangements is comparable with the other species. The extremely low number of large rearrangements in the gorilla genome (dotted green line) is a reflection of the use of the human genome as a template in the assembly process. **c**, Examples of gibbon–human synteny breakpoints. The first two are class I breakpoints (that is, base-pair resolution) originated through non-homology based mechanisms. NLE12_1 is the result of an inversion in human chromosome 1 and NLE18_6 is the result of a translocation between human chromosomes 16 and 5 with an untemplated insertion in the gibbon sequence shown in purple; in both cases, micro-homologies in the human sequences are shown in red. The last example (NLE9_4) is a class II breakpoint (3.2 kb) containing a mixture of repetitive sequences.

(nt)) and/or the absence of identity, suggesting non-homologous end joining. The presence of micro-homologies (2–26 nt) in a small portion of the breakpoints (13/42) pointed to additional alternative mechanisms such as microhomology-mediated end joining[8] or microhomology-mediated break-induced replication[9]. The origin of the complex structure of breakpoint intervals (class II) was less obvious and reinforced the observation that repeats have the tendency to accumulate at the breakpoints.

To explore the possibility that chromatin conformation, rather than sequence, might predispose regions to breakage, we investigated the relationship between gibbon breakpoints and CCCTC-binding factor (CTCF), an evolutionarily conserved protein with multiple functions, including mediating intra- and interchromosomal interactions[10]. We performed chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) of CTCF-bound DNA using lymphoblast cell lines established from eight gibbon individuals (Supplementary Information section S5). We observed an enrichment of gibbon–human breakpoints in CTCF-binding events (P value = 0.0028), which increased when we considered a ~20 kb window centred around each breakpoint (P value of < 0.0001). Notably, this enrichment was maintained only for CTCF-binding events shared with other primates (human, orangutan and rhesus macaque)[11] but not those specific to gibbon (P value = 0.0019) (Supplementary Fig. SF5.4).

Thus, gibbon–human breakpoints co-localized with distinct genomic features and epigenetic marks; however, as many of these features were shared with other primates, other factors unique to the gibbon lineage must have been present to trigger the increased frequency of chromosomal rearrangements.

LAVA insertions in the gibbon genome

The gibbon genome contains all previously described classes of transposable elements that are mostly also present in other primates. One exceptional addition is the LAVA element, a novel retrotransposon that emerged exclusively in gibbons[12] and has a composite structure comprised of portions of other repeats (3′-L1-*Alu*S-VNTR-*Alu*$_{\text{-like}}$-5′) (Fig. 3a). Searches of Nleu1.0 retrieved 1,797 LAVA insertions, 1,256 of which were 3′ intact elements, many carrying signs of target-primed reverse transcription (TPRT)[13]. The distribution of 3′ intact LAVA elements uncovered a significant overlap with genes (Pearson chi-squared, $P = 0.017$) and Gene Ontology (GO) analyses using the database for annotation, visualization, and integrated discovery (DAVID)[14] showed a significant functional enrichment exclusive to the 'microtubule cytoskeleton' category (false discovery rate = 0.031, P value = 0.001) (Supplementary Information section S7 and Supplementary Data File 6) (Extended Data Fig. 3). Additional analyses with meta-pathway database tools[15,16] refined this enrichment to pathways related to chromosome segregation, including 'establishment of sister chromatid cohesion' and 'mitotic metaphase and anaphase' (Supplementary Table ST7.3). Genes with LAVA insertions include proteins that function as checkpoints for cell division and for spindle integrity/architecture (such as *MAP4*, *CEP164* and *BUB1B*)[17–19], participate in kinetochore assembly and attachment to the spindle (for example, *MAD1L1* and *CLASP2*)[20,21], and have a role in chromosome segregation during cell division (for example, *KIFAP3* and *KIF27*)[22] (Extended Data Table 1).

Intragenic LAVA insertions were skewed toward introns (Pearson chi-squared, $P = 0.0001$) and were less frequent than expected when within <1 kb of the nearest exon junction (Extended Data Fig. 3). The majority (74%) of intronic LAVA elements were found in the antisense orientation. We speculated that intronic antisense LAVA insertions may cause early transcription termination by providing a polyadenylation site in the antisense orientation, as previously described for L1 elements[23,24] (Extended Data Fig. 3). Indeed, we found 84.1% of the 3′-intact LAVA elements encoded a perfect polyadenylation signal at their 3′ end in antisense orientation.

To obtain experimental evidence that LAVA elements disrupt transcription, we performed a reporter assay in which the 3' end of a luciferase gene construct lacking a transcriptional termination site was fused to the 3'-terminal fragments of LAVA_E and LAVA_F elements, mimicking the arrangement observed in gibbon genes (Fig. 3b, left). Luciferase activity exceeding background level by ~50% was observed from the LAVA_F reporter construct (Fig. 3b, right), indicating faithful termination of luciferase transcription. Furthermore, 3′ rapid amplification of cDNA ends (RACE) experiments confirmed that the transcription termination site had been supplied from the LAVA element (Extended Data Fig. 3). Thus antisense intronic LAVA insertions can cause early transcription termination with some variability possibly due to the genomic context of the polyadenylation site, which explained the difference between the two reporter constructs.

We also investigated LAVA induced early transcription termination *in vivo* by analyzing RNA-seq data generated for the gibbon named Asia (Supplementary Table ST2.4). Specifically, we looked for paired-end reads only partially aligning to an antisense LAVA element due to untemplated residues and then identified cases for which the presence of a poly(A) tail was preventing full-length alignment. This analysis revealed that elements from a variety of subfamilies have the potential to

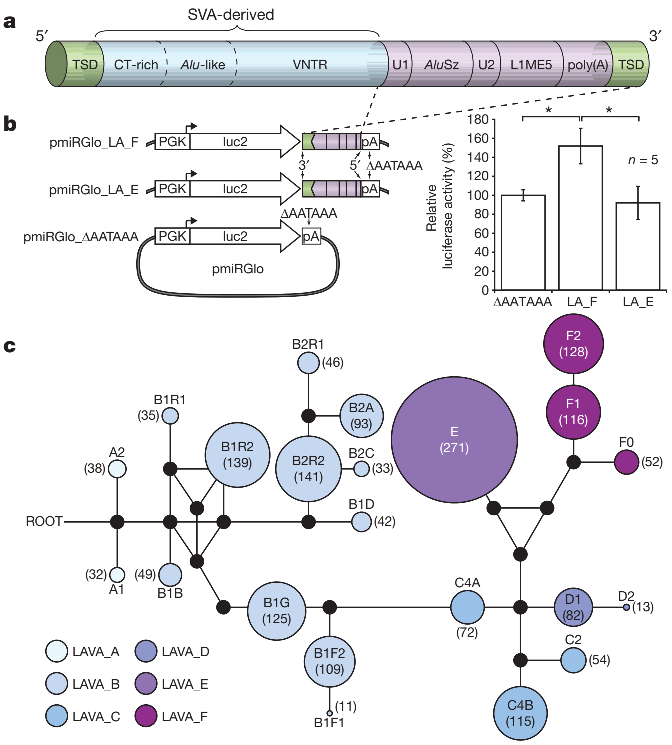


Figure 3 | The LAVA element and evidence for LAVA-mediated early transcription termination. a, Schematic view of the LAVA element highlights the main components that originated from common repeats (L1, *Alu*, VNTR and *Alu*-like). Target-site duplications (TSDs) and the poly(A) tail are also indicated. **b**, Luciferase reporter constructs used to assay for LAVA-mediated early transcriptional termination (left panel) and results of the luciferase reporter assay (right panel) showing increased luciferase activity by ~50% relative to the background for pmiRGlo_LA_F (*$P = 0.0013$) (see Supplementary Information section S7.8) $n = 5$, five biological replicates, from five independent transfections done for each experimental condition tested. The experiment shown was replicated twice in the laboratory. Statistics were carried out using a Student's t-test (two sided), P values for all pairwise comparisons LA_F vs. LA_E, ΔPA vs. LA_F, and ΔPA vs. LA_E respectively (with 95% CI) were adjusted for multiple comparisons according to the Bonferroni method. Centre values show the average, error bars indicate standard deviation. **c**, A median-joining network showing the relationships among the 22 LAVA subfamilies generated by comparing the 3′ intact LAVA elements. Coloured circles represent subfamilies and their size is proportional to the number of elements in the subfamily (numbers inside each circle). Black dots represent hypothetical sequences connecting adjacent subfamilies. All possible relationships are shown. Branch lengths are not drawn to scale.

cause early transcription termination, including those identified for LAVA elements inserted in the microtubule cytoskeleton genes (for example, LAVA_B2R2, LAVA_C4B, LAVA_B1R2) (Extended Data Table 1). We observed that early transcription termination occurred at relatively low levels as we identified a significant number of read pairs indicative of normal transcription and splicing for LAVA-terminated genes (Supplementary Table ST7.5). This is to be expected, as full inactivation of many of these genes would be lethal. On the other hand, as alternative splicing and RNA pol II transcript termination/polyadenylation are tightly coupled processes, LAVA-mediated early transcription termination could also act by differently affecting distinct isoforms and/or influencing the ratio between isoforms. Finally, LAVA insertions may also affect gene expression by functioning as exon traps, as shown for SVA elements[25]. One putative example of an exon trapping event was identified for *HORMAD2*, a gene that monitors the formation of synapsis during crossover[26] (Supplementary Information section S7, Supplementary Table ST7.6, Supplementary Fig. SF7.1–7.2).

As genome reshuffling began in the common ancestor of all extant gibbon species, LAVA insertions must have occurred in key genes before the four genera diverged. We experimentally confirmed the mode and tempo of all 23 LAVA insertions in genes from the microtubule cytoskeleton category using both site-specific PCR and *in silico* methods (Extended Data Figure 4) and found that most of the insertions (15/23) were shared by the four gibbon genera (Supplementary Data File 6). Eleven of the genes match the structural requirements for early transcription termination and five of them are also shared. These genes include *MAP4*, involved in spindle architecture and *CEP164*, a G2/M checkpoint gene whose inactivation results in an aberrant spindle during cell division[18,19] (Extended Data Table 1).

The complex evolutionary history of gibbons

We explored the relationship between LAVA family expansion and evolution of the gibbon lineage and, through analyses of diagnostic mutations, identified 22 LAVA subfamilies (Fig. 3c). In addition, we tested for the presence or absence of 200 LAVA loci from among the evolutionarily youngest elements in each subfamily (Extended Data Fig. 4) across 17 unrelated gibbon individuals and found that 52% of loci were shared among all four genera, whereas 27% were *Nomascus* specific. The remaining LAVA insertions showed a variety of confounding phylogenetic relationships consistent with incomplete lineage sorting (ILS) of ancestral polymorphisms, perhaps as a result of a rapid radiation of gibbon genera (Supplementary Information section S7; Supplementary Table ST7.1–7.2). We used a maximum likelihood method[27] to obtain age estimates for the 22 LAVA subfamilies. In the case of the two oldest subfamilies, LAVA_A1 and LAVA_A2, we obtained estimates of ~18 Myr ago and ~17 Myr ago, respectively (Supplementary Table ST7.3). A coalescent-based methodology implemented in the software G-PhosCS[28] using Nleu1.0 estimated a gibbon–great ape population divergence time of ~16.8 Myr ago (95% confidence intervals (CI): 15.9–17.6 Myr ago) assuming a split time with macaque of 29 Myr ago (Supplementary Information section S4). Hence, the LAVA element probably originated around the time of the divergence of gibbons from the ancestral great ape/human lineage.

The evolutionary history of the gibbon lineage and, in particular, the timing and order of splitting among the four genera, is still a subject of debate[29]. To address this issue, we generated medium coverage (mean ~15×) WGS short read data for two individuals from each of the four genera, including two different *Hylobates* species (*H. moloch* and *H. pileatus*) (Supplementary Table ST2.1–2.2). Although phylogenetic analysis of assembled whole mitochondrial DNA genomes using BEAST[30] strongly supported monophyletic groupings for each gibbon genus, the branching order of the four genera remained unresolved (Supplementary Fig. SF9.1–9.2; Supplementary Information S9).

Neighbour-joining trees constructed from pairwise sequence divergence, *k*, across ~11,000 genic (200 base pairs (bp)) and ~12,000 non-genic (1 kilobase (kb)) autosomal loci supported a supermatrix sequence topology of (((*Siamang* (SSY), *Hoolock* (HLE)), *Nomascus* (NLE)), (*H. pileatus* (HPL)), *H. moloch* (HMO)) (Fig. 4a); nevertheless, bootstrap confidence for the node separating NLE and *Hylobates* was low (~52%). This topology was also the most frequently observed when constructing *k*-based unweighted pair group method with arithmetic mean (UPGMA) trees along the genome using non-overlapping 100-kb sliding windows. However, all 15 possible rooted topologies for the four genera were observed at considerable frequencies (Extended Data Fig. 5), consistent with the extensive ILS observed in the LAVA element analysis.

In order to infer the most likely bifurcating species topology amongst the four genera while taking into account ILS, we used a novel coalescent-based ABC methodology using the autosomal non-genic and genic loci (Veeramah *et al.*, in the press) (Supplementary Information section S8). The topology described above had the highest combined posterior probability, though support was relatively low (*P* (model) = 17%) and other

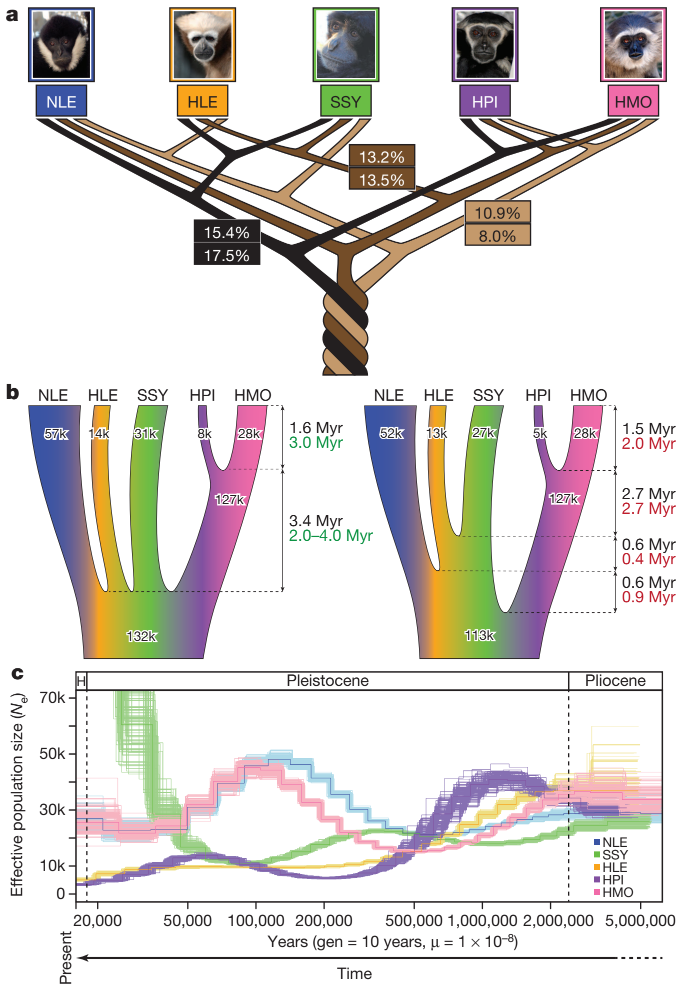


Figure 4 | Gibbon phylogeny and demography. **a**, The three most frequently observed UPGMA gene trees (numbers at the top) constructed across the genome at 100-kb sliding windows and posterior probabilities (numbers at the bottom) for the same species topologies from a coalescent-based ABC analysis. The relatively low numbers observed suggest presence of substantial ILS amongst the gibbon genera. **b**, Parameters estimates describing gibbon population demography assuming an instant radiation for all four genera (left) and the most probable bifurcating species topology (right). Black, green and red numbers indicate divergence times and N_e as calculated by ABC, BEAST and G-PhoCS analysis, respectively (Supplementary Information section S9). **c**, PSMC analysis estimating changes in historical N_e. The large increase in N_e observed in our PSMC plot for SSY in recent times is probably exaggerated due to higher sequencing error and mapping biases in non-NLE samples (see details in Supplementary section S8). A generation time of 10 years[45,46] was used to obtain a per generation mutation rate of 1×10^{-8} per year.

topologies, including one with NLE and *Hylobates* interchanged as the most external taxa, had comparable probabilities (Fig. 4a).

The estimated internal branch lengths under the best species topology using our ABC framework and G-PhoCS were very short, supporting a rapid speciation process for the four gibbon genera (Fig 4b, right). Given this observation and uncertainty in the best topology, we also estimated parameters under an instantaneous speciation model (Fig. 4b, left). Assuming an overall autosomal mutation rate of 1×10^{-9} per site per year, we placed the beginning of the speciation process at ~5 Myr ago under both models, with the two *Hylobates* species diverging ~1.5 Myr ago.

Consistent with the ABC analysis, SSY and HLE share the largest number of alleles across the whole genome (Supplementary Table ST8.5). However, NLE and the two *Hylobates* samples are both significantly closer to SSY than HLE as assessed by the D-statistic[31]. This result could be explained by two independent gene flow events between SSY and both NLE and *Hylobates*. However, fertile intergenic hybrids have yet to be observed either in the wild or captivity[32]; an alternative explanation would be long-term population structure in the gibbon ancestral population. Both the ABC and G-PhoCS analyses suggest that the ancestral gibbon effective population size (N_e) was large (80,000–130,000), but neither of these frameworks can distinguish this from a structured ancestral population.

The coalescent-based analysis (Fig. 4a), along with estimates of genome-wide heterozygosity (Supplementary Fig. ST8.2), suggests a larger long-term N_e for both *N. leucogenys* and *H. moloch* compared to the other species. Analysis using the pairwise sequentially Markovian coalescent (PSMC) model[33] indicates that these two species underwent an increase in N_e during the Late Pleistocene era (500–100 thousand years ago (kyr ago) followed by a subsequent decrease in N_e 100–50 kyr ago (Fig. 4c) (Supplementary Information section S8). Fluctuation in N_e could result from changes in the actual number of individuals in the population, changes in population structure, and/or variable gene flow.

Functional sequence evolution

Accelerated substitution rates are a hallmark of adaptive evolution, and genomic regions with excess lineage-specific substitutions have been found to have functional roles[34]. We identified 240 short (153 bp) median length) regions with accelerated substitution rates in the gibbon lineage (gibARs). We observed that gibARs were primarily intergenic (66%) and tended to co-localize near the same genes as LAVA elements (P value $= 81 \times 10^{-6}$; odds ratio of 2.74 (95% CI: 1.79–4.07)). Consistent with this finding, a GO enrichment test for genes within $\pm$ 100 kb of each gibAR (in comparison with background genes) revealed enrichment for the 'chromosome organization' category (Benjamini–Hochberg false discovery rate <5%) (Extended Data Fig. 6). Given evidence of functional roles gathered for human accelerated regions[35], we speculate that the gibARs may create functional elements (for example, enhancers or protein-binding domains) to modulate the transcriptional effect of local LAVA insertions (Supplementary Information section S12 and Supplementary Data File 9).

We assessed the potential presence of positive selection in 13,638 human genes with one-to-one orthologues in gibbon using a branch-site likelihood ratio test[36] (Supplementary Information section S10). One of the most striking features of gibbons is their use of brachiation (arboreal locomotion using only the arms). We uncovered evidence related to traits possibly associated with this adaptation such as the gibbon's longer arms, more powerful shoulder flexors, rotator muscles and elbow flexors[37]. First, some genes whose functions relate to these anatomical specializations appear to have undergone positive selection in gibbons. They include *TBX5* (P value = 0.00015), required for the development of all forelimb elements[38]; *COL1A1* (pro-alpha1 chains of type I collagen) (P value $= 3.39 \times 10^{-11}$), the fibril-forming collagen that is the main protein of bones, tendons and teeth[39]; and *CHRNA1* (acetylcholine receptor subunit alpha precursor) (P value = 0.00039), involved in skeletal muscle contraction[40]. These genes have not been identified as positively selected in other primates to date. We also observed that some genes involved in chondrogenesis (*SNX19*, *ID2* and *EXT1*) were associated with gibARs. Finally, the chondroadherin gene (*CHAD*)[41] coding for a cartilage matrix protein is specifically duplicated in all gibbon genera (Extended Data Fig. 2).

Discussion

Our sequencing, assembling and analysis of the gibbon genome has provided numerous insights into the accelerated evolution of the gibbon karyotype and identified genetic signatures related to gibbon biology. First, segmental duplications and repetitive sequences were the best predictors of gibbon–human breakpoints, although we excluded a causal role given the predominance of non-homology-based repair signatures. Furthermore, accelerated rearrangement was confined to large-scale chromosomal events, pointing to a mechanism responsible for causing gross chromosomal changes, rather than global genomic instability. This is in line with our hypothesis that the high rate of chromosomal rearrangements may have been due to LAVA-induced premature transcription termination of chromosome segregation genes. This effect may have occurred at a low enough level to be compatible with life but sufficient to increase the frequency of chromosome segregation errors. The link between erroneous chromosome segregation and increased chromosomal rearrangement has been recently demonstrated by others through *in vitro* experiments[25,26].

The question remains how such a high number of chromosomal rearrangements could become fixed in such a relatively short time. One possibility is that a combination of geographic isolation and post-mating reproductive barriers accelerated the radiation of the four gibbon genera. Our estimates dated the lineage-splitting event to the Miocene–Pliocene transition, when major changes in the distribution of tropical and subtropical forests were caused by the elevation of the Yunnan plateau and rise in sea levels[42,43]. Furthermore, fluctuation in sea levels beginning in the Early Pliocene appears to have brought about cycles of forest fragmentation and amalgamation, leading to alternating range compression and expansion for many mammalian groups[44].

Together, these results advance our knowledge of the unique traits of the small apes and highlight the complex evolutionary history of these species. Moreover, our analyses of the rearranged gibbon genome help to provide insight into the mechanisms of chromosome evolution as well as uncovering a new source of genome plasticity.

METHODS SUMMARY

Sanger-based whole-genome sequencing was performed as described for other species. The genome assembly was generated using the ARACHNE genome assembler assisted with alignment data from the human genome (Supplementary Information section S1). The source DNA for the sequencing was derived from a single female (Asia; studbook no. 0098, ISIS no. NLL605) housed at the Virginia Zoo in Norfolk, Virginia. Short-read libraries were constructed at the Oregon Health & Science University (OHSU) following standard Illumina protocols and sequenced on an Illumina HiSeq 2000. Analyses were performed with custom analysis pipelines. See Supplementary Information for additional information about the methods.

Received 23 March; accepted 14 July 2014.

1. Mittermeier, R. A., Rylands, A. B. & Wilson, D. E. *Handbook of the Mammals of the World* Vol. 3 (Lynx Edicions, (2013).
2. Carbone, L. *et al.* A high-resolution map of synteny disruptions in gibbon and human genomes. *PLoS Genet.* **2,** e223 (2006).
3. Locke, D. P. *et al.* Comparative and demographic analysis of orang-utan genomes. *Nature* **469,** 529–533 (2011).
4. Gibbs, R. A. *et al.* Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316,** 222–234 (2007).
5. Girirajan, S. *et al.* Sequencing human–gibbon breakpoints of synteny reveals mosaic new insertions at rearrangement sites. *Genome Res.* **19,** 178–190 (2009).
6. Carbone, L. *et al.* Evolutionary breakpoints in the gibbon suggest association between cytosine methylation and karyotype evolution. *PLoS Genet.* **5,** e1000538 (2009).
7. Bailey, J. A. & Eichler, E. E. Primate segmental duplications: crucibles of evolution, diversity and disease. *Nature Rev. Genet.* **7,** 552–564 (2006).
8. Yan, C. T. *et al.* IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* **449,** 478–482 (2007).

9. Hastings, P. J., Ira, G. & Lupski, J. R. A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet.* **5,** e1000327 (2009).
10. Merkenschlager, M. & Odom, D. T. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell* **152,** 1285–1297 (2013).
11. Schwalie, P. C. *et al.* Co-binding by YY1 identifies the transcriptionally active, highly conserved set of CTCF-bound regions in primate genomes. *Genome Biol.* **14,** R148 (2013).
12. Carbone, L. *et al.* Centromere remodeling in *Hoolock leuconedys* (Hylobatidae) by a new transposable element unique to the gibbons. *Genome Biol. Evol.* **4,** 648–658 (2012).
13. Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72,** 595–605 (1993).
14. Huang da W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4,** 44–57 (2009).
15. Mostafavi, S., Ray, D., Warde-Farley, D., Grouios, C. & Morris, Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol.* **9** (Suppl. 1), S4 (2008).
16. Kamburov, A., Wierling, C., Lehrach, H. & Herwig, R. ConsensusPathDB—a database for integrating human functional interaction networks. *Nucleic Acids Res.* **37,** D623–D628 (2009).
17. Baker, D. J., Jin, F., Jeganathan, K. B. & van Deursen, J. M. Whole chromosome instability caused by Bub1 insufficiency drives tumorigenesis through tumor suppressor gene loss of heterozygosity. *Cancer Cell* **16,** 475–486 (2009).
18. Samora, C. P. *et al.* MAP4 and CLASP1 operate as a safety mechanism to maintain a stable spindle position in mitosis. *Nature Cell Biol.* **13,** 1040–1050 (2011).
19. Leber, B. *et al.* Proteins required for centrosome clustering in cancer cells. *Sci. Transl. Med.* **2,** 33ra38 (2010).
20. Schuyler, S. C., Wu, Y. F. & Kuan, V. J. The Mad1–Mad2 balancing act—a damaged spindle checkpoint in chromosome instability and cancer. *J. Cell Sci.* **125,** 4197–4206 (2012).
21. Maia, A. R. *et al.* Cdk1 and Plk1 mediate a CLASP2 phospho-switch that stabilizes kinetochore-microtubule attachments. *J. Cell Biol.* **199,** 285–301 (2012).
22. Haraguchi, K., Hayashi, T., Jimbo, T., Yamamoto, T. & Akiyama, T. Role of the kinesin-2 family protein, KIF3, during mitosis. *J. Biol. Chem.* **281,** 4094–4099 (2006).
23. Han, J. S., Szak, S. T. & Boeke, J. D. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature* **429,** 268–274 (2004).
24. Wheelan, S. J., Aizawa, Y., Han, J. S. & Boeke, J. D. Gene-breaking: a new paradigm for human retrotransposon-mediated gene evolution. *Genome Res.* **15,** 1073–1078 (2005).
25. Damert, A. *et al.* 5′-Transducing SVA retrotransposon groups spread efficiently throughout the human genome. *Genome Res.* **19,** 1992–2008 (2009).
26. Wojtasz, L. *et al.* Meiotic DNA double-strand breaks and chromosome asynapsis in mice are monitored by distinct HORMAD2-independent and -dependent mechanisms. *Genes Dev.* **26,** 958–973 (2012).
27. Marchani, E. E., Xing, J., Witherspoon, D. J., Jorde, L. B. & Rogers, A. R. Estimating the age of retrotransposon subfamilies using maximum likelihood. *Genomics* **94,** 78–82 (2009).
28. Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G. & Siepel, A. Bayesian inference of ancient human demography from individual genome sequences. *Nature Genet.* **43,** 1031–1034 (2011).
29. Wall, J. D. *et al.* Incomplete lineage sorting is common in extant gibbon genera. *PLoS ONE* **8,** e53682 (2013).
30. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7,** 214 (2007).
31. Durand, E. Y., Patterson, N., Reich, D. & Slatkin, M. Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* **28,** 2239–2252 (2011).
32. Hirai, H., Hirai, Y., Domae, H. & Kirihara, Y. A most distant intergeneric hybrid offspring (Larcon) of lesser apes, *Nomascus leucogenys* and *Hylobates lar. Hum. Genet.* **122,** 477–483 (2007).
33. Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* **475,** 493–496 (2011).
34. Prabhakar, S. *et al.* Human-specific gain of function in a developmental enhancer. *Science* **321,** 1346–1350 (2008).
35. Pollard, K. S. *et al.* An RNA gene expressed during cortical development evolved rapidly in humans. *Nature* **443,** 167–172 (2006).
36. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24,** 1586–1591 (2007).
37. Michilsens, F., Vereecke, E. E., D'Août, K. & Aerts, P. Functional anatomy of the gibbon forelimb: adaptations to a brachiating lifestyle. *J. Anat.* **215,** 335–354 (2009).
38. Browne, M. L. *et al.* Evaluation of genes involved in limb development, angiogenesis, and coagulation as risk factors for congenital limb deficiencies. *Am. J. Med. Genet. A.* **158A,** 2463–2472 (2012).
39. Marini, J. C. *et al.* Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans. *Hum. Mutat.* **28,** 209–221 (2007).
40. Masuda, A. *et al.* hnRNP H enhances skipping of a nonfunctional exon P3A in *CHRNA1* and a mutation disrupting its binding causes congenital myasthenic syndrome. *Hum. Mol. Genet.* **17,** 4022–4035 (2008).
41. Hessle, L. *et al.* The skeletal phenotype of chondroadherin deficient mice. *PLoS ONE* **8,** e63080 (2013).
42. Cane, M. A. & Molnar, P. Closing of the Indonesian seaway as a precursor to east African aridification around 3–4 million years ago. *Nature* **411,** 157–162 (2001).
43. Xu J.-X., Ferguson D. K., Li, C.-S. & Wang Y.-F. Late Miocene vegetation and climate of the Lühe region in Yunnan, southwestern China. *Rev. Palaeobot. Palynol.* **148,** 36–59 (2008).
44. Woodruff D. S. & Turner L. M.. The Indochinese–Sundaic zoogeographic transition: a description and analysis of terrestrial mammal species distributions. *J. Biogeogr.* **36,** 803–821 (2009).
45. Harvey, P. H., Martin, R. D. & Clutton-Brock, T. H. in *Primate Societies* (eds Smuts B. B., *et al.*) Life histories in comparative perspective. 181–196 (Chicago Univ. Press, 1987).
46. Kim, S. K. *et al.* Patterns of genetic variation within and between Gibbon species. *Mol. Biol. Evol.* **28,** 2211–2218 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements The gibbon genome project was funded by the National Human Genome Research Institute (NHGRI) including grants U54 HG003273 (R.A.G.) and U54 HG003079 (R.K.W.) with further support from National Institutes of Health NIH/NIAAA P30 AA019355 and NIH/NCRR P51 RR000163 (L.C.), R01_HG005226 (J.D.W., M.F.H.), NIH P30CA006973 (S.J.W.), a fellowship from the National Library of Medicine Biomedical Informatics Research Training Program (N.H.L.), R01 GM59290 (M.A.B.) and U41 HG007497-01 (M.A.B, M.K.K.), R01 MH081203 (J.M.S.), HG002385 (E.E.E.), National Science Foundation (NSF) CNS-1126739 (B.U., M.A.B., M.K.K.) and DBI-0845494 (M.W.H.), PRIN 2012 (M.R.), Futuro in ricerca 2010 RBFR103CE3 (M.V.), ERC Starting Grant (260372) and MICINN (Spain) BFU2011-28549 (T.M.-B.), grant of the Ministry of National Education, CNCS – UEFISCDI, project number PN-II-ID-PCE-2012-4-0090 (A.D.), grant of the Deutsche Forschungsgemeinschaft SCHU1014/8-1 (G.G.S.), ERC Starting and Advanced Grant and EMBO Young Investigator Award (Z.I., N.V.F.), ERC Starting Grant and EMBO Young Investigator Award (D.T.O.), Commonwealth Scholarship Commission (M.C.W.). E.E.E. is an investigator of the Howard Hughes Medical Institute. We acknowledge the contributions of the staff of the HGSC, including the operations team: H. Dinh, S. Jhangiani V. Korchina, C. Kovar; the library team: K. Blankenburg, L. Pu, S. Vattathil; the assembly team: D. Rio-Deiros, H. Jiang; the submissions team: M. Batterton, D. Kalra, K. Wilczek-Boney, W. Hale, G. Fowler, J. Zhang; the quality control team: P. Aqrawi, S. Gross, V. Joshi, J. Santibanez; and the sequence production team: U. Anosike, C. Babu, D. Bandaranaike, B. Beltran, D. Berhane-Mersha, C. Bickham, T. Bolden, M. Dao, M. Davila, L. Davy-Carroll, S. Denson, P. Fernando, C. Francis, R. Garcia III, B. Hollins, B. Johnson, J. Jones, J. Kalu, N. Khan, B. Leal, F. Legall III, Y. Liu, J. Lopez, R. Mata, M. Obregon, C. Onwere, A. Parra, Y. Perez, A. Perez, C. Pham, J. Quiroz, S. Ruiz, M. Scheel, D. Simmons, I. Sisson, J. Tisius, G. Toledanes, R. Varghese, V. Vee, D. Walker, C. White, A. Williams, R. Wright, T. Attaway, T. Garrett, C. Mercado, N. Ngyen, H. Paul and Z. Trejos. We thank Z. Ivics for providing some of the reagents. We additionally acknowledge the Production Sequencing Group at The Genome Institute. Wellcome Trust (grant numbers WT095908 and WT098051), NHGRI (U41HG007234) and European Molecular Biology Laboratory. For the production of next-generation sequences, we acknowledge the Massively Parallel Sequencing Shared Resources (MPSSR) at OHSU, the National Center of Genomic Analyses (CNAG) (Barcelona, Spain), the University of Arizona Genetics Core (UAGC), and the UCSF sequencing core. We also acknowledge the Louisiana Optical Network Institute (LONI). We thank the Gibbon Conservation Center and the Fort Wayne Children's Zoo for providing the gibbon samples. The MAKER annotation pipeline is supported by NSF IOS-1126998.We thank T. Brown for proofreading and editing the manuscript.

Author Contributions L.C. led the project and the manuscript preparation. L.C., W.C.W., K.C.W., J.R., E.E.E., T.M.-B., R.A.H., K.R.V. and M.F.H. supervised the project and contributed to overall organization of the manuscript. L.C. and T.J.M. prepared the figures. Sanger data production, assembly construction and testing was carried out by: L.F., C.F., D.M.M., L.V.N., A.C., S.L.L., L.R.L., D.P.L., W.C.W., K.C.W., J.R., S.G., L.D.W.H., D.R. and S.M. Mitochondrial genome assembly was done by Y.L. Illumina sequencing production and submission: L.C., T.M.-B., J.D.W., M.F.H., E.T., L.J.W., M.G., I.G., A.B. and J.H.-R. Samples were provided by G.S. Gene set and validation of gene models: D.B., S.W., S.S., B.A., M.M., J.He., P.F., M.S.C. and M.Y. Assembly validation: B.L.-G., J.He. and T.M.-B. BAC library generation: P.J.dJ., B.tH. and B.Z. Cytogenetic analyses: M.R., N.A. and O.C. Segmental duplications and structural variations: J.Hu., C.B., B.L.-G., J.Q., M.F.-C., G.C., F.A., M.V., T.M.-B. and E.E.E. cDNA Array CGH: L.D., M.O'B., A.K.-F. and J.M.S. Comparative analysis of gibbon chromosomal rearrangements was carried out by J.He. Breakpoint analysis: L.C., C.W.W. and L.J.W. LAVA analysis: L.C., R.A.H., T.J.M., N.H.L., L.J.W., K.A.N., K.S., A.D., M.A.B., M.K.K., J.A.W., B.U., A.S. and R.H. Luciferase assay and 3′ RACE: A.D., B.I., C.O., G.G.S., N.V.F. and Z.I. RNA-seq analysis for early transcription termination: S.J.W. and C.L.B. Short-read alignments, SNP calling and population genetics analysis (autosomal DNA): L.M.J., F.L.M., A.E.W., L.J.W., K.R.V., M.F.H. and J.D.W. Population genetics analyses (mtDNA): C.R., L.W., M.B. and T.M.-B. Positive selection analyses: G.W.C.T. and M.W.H. Gene family evolution analyses: M.W.H. and C.C. Gibbon accelerated region analyses: K.S.P. and D.K. CTCF-binding analyses: M.C.W., D.T.O., P.F., E.T., C.W.W., L.J.W., J.He. and K.B. Biogeography analysis: N.G.J. and C.R. Principal investigators: R.K.W. and R.A.G.

Author Information The *N. leucogenys* WGS project has been deposited in GenBank under the project accession ADFV00000000.1. All short-read data have been deposited into the Short Read Archive (http://www.ncbi.nlm.nih.gov/sra) under the accession number SRP043117. Resources for exploring the gibbon genome are available at UCSC (http://genome.ucsc.edu), Ensembl (http://ensembl.org), NCBI (http://ncbi.nlm.nih.gov), and the Baylor College of Medicine Human Genome Sequencing Center (https://www.hgsc.bcm.edu/non-human-primates/gibbon-genome-project). This paper is dedicated to the memory of Alan R. Mootnick (1951–2011). Reprints and permissions information is available at

www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.C. (carbone@ohsu.edu).

Lucia Carbone[1,2,3,4], R. Alan Harris[5], Sante Gnerre[6], Krishna R. Veeramah[7,8], Belen Lorente-Galdos[9], John Huddleston[10,11], Thomas J. Meyer[1], Javier Herrero[12,13]†, Christian Roos[14], Bronwen Aken[12,15], Fabio Anaclerio[16], Nicoletta Archidiacono[16], Carl Baker[10], Daniel Barrell[12,15], Mark A. Batzer[17], Kathryn Beal[12], Antoine Blancher[18], Craig L. Bohrson[19], Markus Brameier[14], Michael S. Campbell[20], Oronzo Capozzi[16], Claudio Casola[21], Giorgia Chiatante[16], Andrew Cree[22], Annette Damert[23], Pieter J. de Jong[24], Laura Dumas[25], Marcos Fernandez-Callejo[9], Paul Flicek[12], Nina V. Fuchs[26], Ivo Gut[27], Marta Gut[27], Matthew W. Hahn[28], Jessica Hernandez-Rodriguez[9], LaDeana W. Hillier[29], Robert Hubley[30], Bianca Ianc[23], Zsuzsanna Izsvák[26], Nina G. Jablonski[31], Laurel M. Johnstone[7], Anis Karimpour-Fard[25], Miriam K. Konkel[17], Dennis Kostka[32], Nathan H. Lazar[4], Sandra L. Lee[22], Lora R. Lewis[22], Yue Liu[22], Devin P. Locke[29]†, Swapan Mallick[33], Fernando L. Mendez[7]†, Matthieu Muffato[12], Lynne V. Nazareth[22], Kimberly A. Nevonen[2], Majesta O'Bleness[25], Cornelia Ochis[23], Duncan T. Odom[15,34], Katherine S. Pollard[35,36,37], Javier Quilez[9], David Reich[33], Mariano Rocchi[16], Gerald G. Schumann[38], Stephen Searle[15], James M. Sikela[25], Gabriella Skollar[39], Arian Smit[29], Kemal Sonmez[4,40], Boudewijn ten Hallers[24]†, Elizabeth Terhune[2], Gregg W. C. Thomas[28], Brygg Ullmer[41], Mario Ventura[16], Jerilyn A. Walker[17], Jeffrey D. Wall[36,37], Lutz Walter[14], Michelle C. Ward[34]†, Sarah J. Wheelan[19], Christopher W. Whelan[40]†, Simon White[15], Larry J. Wilhelm[2], August E. Woerner[7], Mark Yandell[20,42], Baoli Zhu[24]†, Michael F. Hammer[7], Tomas Marques-Bonet[9,27], Evan E. Eichler[10,11], Lucinda Fulton[29], Catrina Fronick[29], Donna M. Muzny[22], Wesley C. Warren[29], Kim C. Worley[22], Jeffrey Rogers[22], Richard K. Wilson[29] & Richard A. Gibbs[22]

[1]Oregon Health & Science University, Department of Behavioral Neuroscience, 3181 SW Sam Jackson Park Road Portland, Oregon 97239, USA. [2]Oregon National Primate Research Center, Division of Neuroscience, 505 NW 185th Avenue, Beaverton, Oregon 97006, USA. [3]Oregon Health & Science University, Department of Molecular & Medical Genetics, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239, USA. [4]Oregon Health & Science University, Bioinformatics and Computational Biology Division, Department of Medical Informatics & Clinical Epidemiology, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239, USA. [5]Baylor College of Medicine, Department of Molecular and Human Genetics, One Baylor Plaza, Houston, Texas 77030, USA. [6]Nabsys, 60 Clifford Street, Providence, Rhode Island 02903, USA. [7]University of Arizona, ARL Division of Biotechnology, Tucson, Arizona 85721, USA. [8]Stony Brook University, Department of Ecology and Evolution, Stony Brook, New York 11790, USA. [9]IBE, Institut de Biologia Evolutiva (UPF-CSIC), Universitat Pompeu Fabra, PRBB, Doctor Aiguader, 88, 08003 Barcelona, Spain. [10]Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington 98195, USA. [11]Howard Hughes Medical Institute, 1705 NE Pacific Street, Seattle, Washington 98195, USA. [12]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK. [13]The Genome Analysis Centre, Norwich Research Park, Norwich NR4 7UH, UK. [14]Leibniz Institute for Primate Research, Gene Bank of Primates, German Primate Center, Göttingen 37077, Germany. [15]European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK. [16]University of Bari, Department of Biology, Via Orabona 4, 70125, Bari, Italy. [17]Louisiana State University, Department of Biological Sciences, Baton Rouge, Louisiana 70803, USA. [18]University of Paul Sabatier, Toulouse 31062, France. [19]The Johns Hopkins University School of Medicine, Department of Oncology, Division of Biostatistics and Bioinformatics, Baltimore, Maryland 21205, USA. [20]University of Utah, Salt Lake City, Utah 84112, USA. [21]Texas A&M University, Department of Ecosystem Science and Management, College Station, Texas 77843, USA. [22]Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA. [23]Babes-Bolyai-University, Institute for Interdisciplinary Research in Bio-Nano-Sciences, Molecular Biology Center, Cluj-Napoca 400084, Romania. [24]Children's Hospital Oakland Research Institute, BACPAC Resources, Oakland, California 94609, USA. [25]University of Colorado School of Medicine, Department of Biochemistry and Molecular Genetics, Aurora, Colorado 80045, USA. [26]Max Delbrück Center for Molecular Medicine, Berlin 13125, Germany. [27]Centro Nacional de Análisis Genómico (CNAG), Parc Científic de Barcelona, Barcelona 08028, Spain. [28]Indiana University, School of Informatics and Computing, Bloomington, Indiana 47408, USA. [29]The Genome Center at Washington University, Washington University School of Medicine, 4444 Forest Park Avenue, Saint Louis, Missouri 63108, USA. [30]Institute for Systems Biology, Seattle, Washington 98109-5234, USA. [31]The Pennsylvania State University, Department of Anthropology, University Park, Pennsylvania 16802, USA. [32]University of Pittsburgh School of Medicine, Department of Developmental Biology, Department of Computational and Systems Biology, Pittsburg, Pennsylvania 15261, USA. [33]Harvard Medical School, Department of Genetics, Boston, Massachusetts 02115, USA. [34]University of Cambridge, Cancer Research UK-Cambridge Institute, Cambridge CB2 0RE, UK. [35]University of California, Gladstone Institutes, San Francisco, California 94158-226, USA. [36]Institute for Human Genetics, University of California, San Francisco, California 94143-0794, USA. [37]Division of Biostatistics, University of California, San Francisco, California 94143-0794, USA. [38]Paul Ehrlich Institute, Division of Medical Biotechnology, 63225 Langen, Germany. [39]Gibbon Conservation Center, 19100 Esguerra Rd, Santa Clarita, California 91350, USA. [40]Oregon Health & Science University, Center for Spoken Language Understanding, Institute on Development and Disability, Portland, Oregon 97239, USA. [41]Louisiana State University, School of Electrical Engineering and Computer Science, Baton Rouge, Louisiana 70803, USA. [42] USTAR Center for Genetic Discovery, University of Utah, Salt Lake City, Utah 84112, USA. †Present addresses: Bill Lyons Informatics Center, UCL Cancer Institute, University College London, London WC1E 6DD, UK (J.He); Seven Bridges Genomics, Cambridge, Massachusetts 02138, USA (D.P.L.); Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA (F.L.M.); BioNano Genomics, San Diego, California 92121, USA (B.t.H.); University of Chicago, Department of Human Genetics, Chicago, Illinois 60637, USA (M.C.W.); Stanley Center for Psychiatric Research, Broad Institute, Cambridge, Massachusetts 02138, USA (C.W.W.); The CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China (B.Z.).

doi:10.1038/nature13480

Comprehensive molecular characterization of gastric adenocarcinoma

The Cancer Genome Atlas Research Network*

Gastric cancer is a leading cause of cancer deaths, but analysis of its molecular and clinical characteristics has been complicated by histological and aetiological heterogeneity. Here we describe a comprehensive molecular evaluation of 295 primary gastric adenocarcinomas as part of The Cancer Genome Atlas (TCGA) project. We propose a molecular classification dividing gastric cancer into four subtypes: tumours positive for Epstein–Barr virus, which display recurrent *PIK3CA* mutations, extreme DNA hypermethylation, and amplification of *JAK2*, *CD274* (also known as *PD-L1*) and *PDCD1LG2* (also known as *PD-L2*); microsatellite unstable tumours, which show elevated mutation rates, including mutations of genes encoding targetable oncogenic signalling proteins; genomically stable tumours, which are enriched for the diffuse histological variant and mutations of *RHOA* or fusions involving RHO-family GTPase-activating proteins; and tumours with chromosomal instability, which show marked aneuploidy and focal amplification of receptor tyrosine kinases. Identification of these subtypes provides a roadmap for patient stratification and trials of targeted therapies.

Gastric cancer was the world's third leading cause of cancer mortality in 2012, responsible for 723,000 deaths[1]. The vast majority of gastric cancers are adenocarcinomas, which can be further subdivided into intestinal and diffuse types according to the Lauren classification[2]. An alternative system, proposed by the World Health Organization, divides gastric cancer into papillary, tubular, mucinous (colloid) and poorly cohesive carcinomas[3]. These classification systems have little clinical utility, making the development of robust classifiers that can guide patient therapy an urgent priority.

The majority of gastric cancers are associated with infectious agents, including the bacterium *Helicobacter pylori*[4] and Epstein–Barr virus (EBV). The distribution of histological subtypes of gastric cancer and the frequencies of *H. pylori* and EBV associated gastric cancer vary across the globe[5]. A small minority of gastric cancer cases are associated with germline mutation in E-cadherin (*CDH1*)[6] or mismatch repair genes[7] (Lynch syndrome), whereas sporadic mismatch repair-deficient gastric cancers have epigenetic silencing of *MLH1* in the context of a CpG island methylator phenotype (CIMP)[8]. Molecular profiling of gastric cancer has been performed using gene expression or DNA sequencing[9–12], but has not led to a clear biologic classification scheme. The goals of this study by The Cancer Genome Atlas (TCGA) were to develop a robust molecular classification of gastric cancer and to identify dysregulated pathways and candidate drivers of distinct classes of gastric cancer.

Sample set and molecular classification

We obtained gastric adenocarcinoma primary tumour tissue (fresh frozen) from 295 patients not treated with prior chemotherapy or radiotherapy (Supplementary Methods S1). All patients provided informed consent, and local Institutional Review Boards approved tissue collection. We used germline DNA from blood or non-malignant gastric mucosa as a reference for detecting somatic alterations. Non-malignant gastric samples were also collected for DNA methylation ($n = 27$) and expression ($n = 29$) analyses. We characterized samples using six molecular platforms (Supplementary Methods S2–S7): array-based somatic copy number analysis, whole-exome sequencing, array-based DNA methylation profiling, messenger RNA sequencing, microRNA (miRNA) sequencing and reverse-phase protein array (RPPA), with 77% of the tumours tested by all six platforms. Microsatellite instability (MSI) testing was performed on all tumour DNA, and low-pass (~6× coverage) whole genome sequencing on 107 tumour/germline pairs.

To define molecular subgroups of gastric cancer we first performed unsupervised clustering on data from each molecular platform (Supplementary Methods S2–S7) and integrated these results, yielding four groups (Supplementary Methods S10.2). The first group of tumours was significantly enriched for high EBV burden ($P = 1.5 \times 10^{-18}$) and showed extensive DNA promoter hypermethylation. A second group was enriched for MSI ($P = 2.1 \times 10^{-32}$) and showed elevated mutation rates and hypermethylation (including hypermethylation at the *MLH1* promoter). The remaining two groups were distinguished by the presence or absence of extensive somatic copy-number aberrations (SCNAs). As an alternative means to define distinct gastric cancer subgroups, we performed integrative clustering of multiple data types using iCluster[13] (Supplementary Methods S10.3). This analysis again indicated that EBV, MSI and the level of SCNAs characterize distinct subgroups (Supplementary Fig. 10.3). Based upon these results from analysis of all molecular platforms, we created a decision tree to categorize the 295 gastric cancer samples into four subtypes (Fig. 1a, b) using an approach that could more readily be applied to gastric cancer tumours in clinical care. Tumours were first categorized by EBV-positivity (9%), then by MSI-high status, hereafter called MSI (22%), and the remaining tumours were distinguished by degree of aneuploidy into those termed genomically stable (20%) or those exhibiting chromosomal instability (CIN; 50%).

Evaluation of the clinical and histological characteristics of these molecular subtypes revealed enrichment of the diffuse histological subtype in the genomically stable group ($40/55 = 73\%$, $P = 7.5 \times 10^{-17}$) (Fig. 1c), an association not attributable to reduced SCNA detection in low purity tumours (Supplementary Fig. 2.8). Each subtype was found throughout the stomach, but CIN tumours showed elevated frequency in the gastroesophageal junction/cardia (65%, $P = 0.012$), whereas most EBV-positive tumours were present in the gastric fundus or body (62%, $P = 0.03$). Genomically stable tumours were diagnosed at an earlier age (median age 59 years, $P = 4 \times 10^{-7}$), whereas MSI tumours were diagnosed at relatively older ages (median 72 years, $P = 5 \times 10^{-5}$). MSI patients tended to be female (56%, $P = 0.001$), but most EBV-positive

*A list of authors and affiliations appears at the end of the paper.

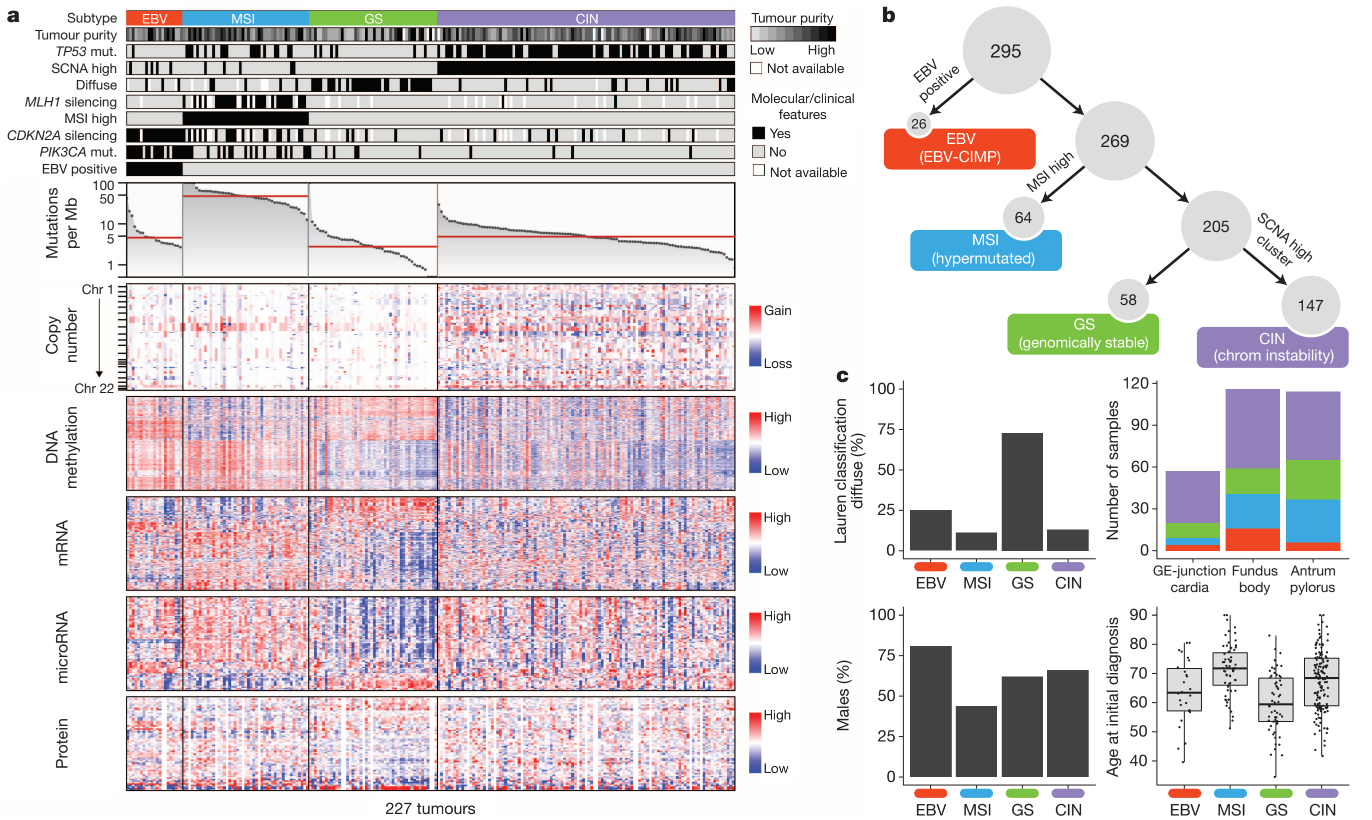


Figure 1 | Molecular subtypes of gastric cancer. **a**, Gastric cancer cases are divided into subtypes: Epstein–Barr virus (EBV)-positive (red), microsatellite instability (MSI, blue), genomically stable (GS, green) and chromosomal instability (CIN, light purple) and ordered by mutation rate. Clinical (top) and molecular data (top and bottom) from 227 tumours profiled with all six platforms are depicted. **b**, A flowchart outlines how tumours were classified into molecular subtypes. **c**, Differences in clinical and histological characteristics among subtypes with subtypes coloured as in **a**, **b**. The plot of patient age at initial diagnosis shows the median, 25th and 75th percentile values (horizontal bar, bottom and top bounds of the box), and the highest and lowest values within 1.5 times the interquartile range (top and bottom whiskers, respectively). GE, gastroesophageal.

cases were male (81%, $P = 0.037$), as previously reported[14]. We did not observe any systematic differences in distribution of subtypes between patients of East Asian and Western origin (Supplementary Methods S1.8). Initial outcome data from this cohort did not reveal survival differences between the four subgroups (Supplementary Information S1.7).

EBV-associated DNA hypermethylation

EBV is found within malignant epithelial cells in 9% of gastric cancers[14]. EBV status was determined using mRNA, miRNA, exome and whole-genome sequencing, yielding highly concordant results (Supplementary Fig. 9.7). By contrast, we detected only sporadic evidence of *H. pylori*, which may reflect the decline of bacterial counts accompanying the progression from chronic gastritis to subsequent carcinoma, as well as technical loss of luminal bacteria during specimen processing. Unsupervised clustering of CpG methylation performed on unpaired tumour samples revealed that all EBV-positive tumours clustered together and exhibited extreme CIMP, distinct from that in the MSI subtype[8], consistent with prior reports[15] (Fig. 2a). Differences between the EBV-CIMP and MSI-associated gastric-CIMP methylation profiles of tumours mirrored differences between these groups in their spectra of mutations (Fig. 1a) and gene expression (Supplementary Fig. 10.6a). EBV-positive tumours had a higher prevalence of DNA hypermethylation than any cancers reported by TCGA (Supplementary Fig. 4.6). All EBV-positive tumours assayed displayed *CDKN2A* (*p16^{INK4A}*) promoter hypermethylation, but lacked the *MLH1* hypermethylation characteristic of MSI-associated CIMP[16]. Genes with promoter hypermethylation most differentially silenced in EBV-positive gastric cancer are shown in Supplementary Table 4.3.

We observed strong predilection for *PIK3CA* mutation in EBV-positive gastric cancer as suggested by prior reports[17,18], with non-silent *PIK3CA* mutations found in 80% of this subgroup ($P = 9 \times 10^{-12}$), including 68% of cases with mutations at sites recurrent in this data set or in the COSMIC repository. In contrast, 3 to 42% of tumours in the other subtypes displayed *PIK3CA* mutations. PI(3)-kinase inhibition therefore warrants evaluation in EBV-positive gastric cancer. *PIK3CA* mutations were more dispersed in EBV-positive cancers, but localized in the kinase domain (exon 20) in EBV-negative cancers (Fig. 2b). The most highly transcribed EBV viral mRNAs and miRNAs fell within the BamH1A region of the viral genome (Supplementary Fig. 9.8) and showed similar expression patterns across tumours, as reported separately[19].

Somatic genomic alterations

To identify recurrently mutated genes, we analysed the 215 tumours with mutation rates below 11.4 mutations per megabase (Mb) (none of which were MSI-positive) separately from the 74 'hypermutated' tumours. Within the hypermutated tumours, we excluded from analysis 11 cases with a distinctly higher mutational burden above 67.7 mutations per Mb (including one tumour with an inactivating *POLE* mutation[20,21]) (Supplementary Information S3.2–3.3), because their large numbers of mutations unduly influence analysis. We used the MutSigCV[22] tool to define recurrent mutations in the 63 remaining hypermutated tumours by first evaluating only base substitution mutations, identifying 10 significantly mutated genes, including *TP53, KRAS, ARID1A, PIK3CA, ERBB3, PTEN* and *HLA-B* (Supplementary Table 3.5). We found *ERBB3* mutations in 16 of 63 tumours, with 13 of these tumours having mutations at recurrent sites or sites reported in COSMIC. MutSigCV analysis including insertions/deletions expanded the list of statistically significant mutated genes to 37, including *RNF43, B2M* and *NF1* (Supplementary Fig. 3.9). Similarly, HotNet analysis of genes mutated within MSI tumours revealed common alterations in major histocompatibility

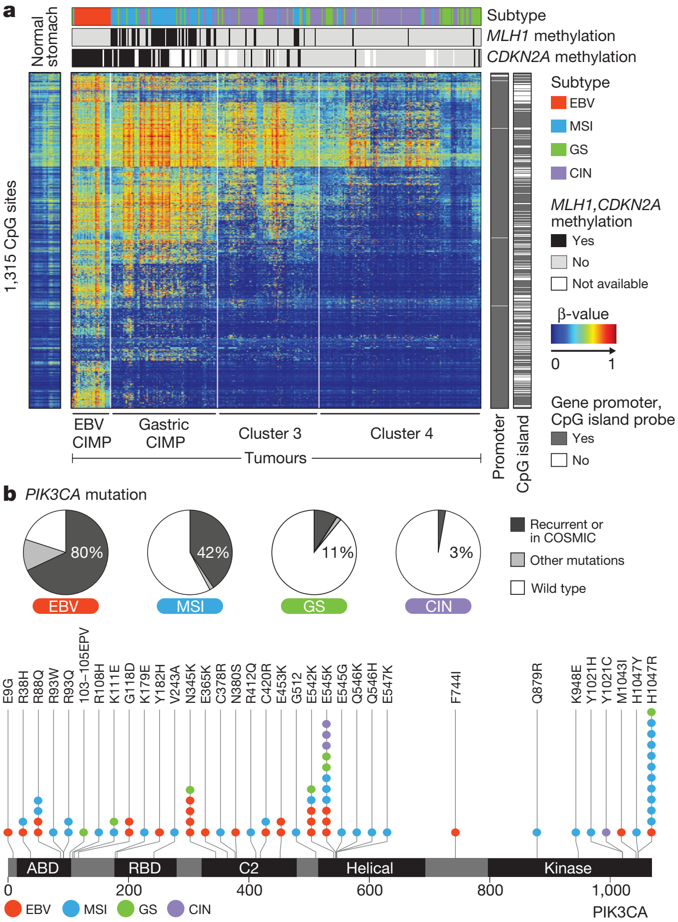


Figure 2 | Molecular characteristics of EBV-positive gastric cancers. a, The heatmap represents unsupervised clustering of DNA methylation at CpG sites for 295 tumours into four clusters: EBV-CIMP ($n = 28$), Gastric-CIMP ($n = 77$), cluster 3 ($n = 73$) and cluster 4 ($n = 117$). Profiles for non-malignant gastric mucosa are to the left of the tumours. **b**, The proportion of tumours harbouring *PIK3CA* mutation in the molecular subtypes with mutations at sites noted recurrently in this data set or in the COSMIC database marked separately (top). Locations of *PIK3CA* mutations with the subtype of the sample with each mutation colour-coded (bottom).

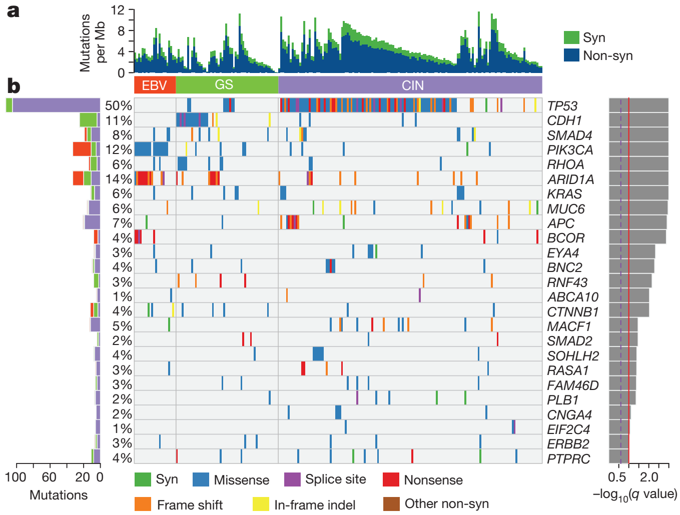


Figure 3 | Significantly mutated genes in non-hypermutated gastric cancer. a, Bars represent somatic mutation rate for the 215 samples with synonymous and non-synonymous mutation rates distinguished by colour. **b**, Significantly mutated genes, identified by MutSigCV, are ranked by the q value (right) with samples grouped by subtype. Mutation colour indicates the class of mutation.

complex class I genes, including *B2M* and *HLA-B* (Supplementary Fig. 11.5–11.7). *B2M* mutations in colorectal cancers and melanoma result in loss of expression of HLA class 1 complexes[23], suggesting these events benefit hypermutated tumours by reducing antigen presentation to the immune system.

Through MutSigCV analysis of the 215 non-hypermutated tumours, we identified 25 significantly mutated genes (Fig. 3). This gene list again included *TP53, ARID1A, KRAS, PIK3CA* and *RNF43*, but also genes in the β-catenin pathway (*APC* and *CTNNB1*), the TGF-β pathway (*SMAD4* and *SMAD2*), and *RASA1*, a negative regulator of RAS. *ERBB2*, a therapeutic target, was significantly mutated, with 10 of 15 mutations occurring at known hotspots; four cases had the S310F *ERBB2* mutation that is activating and drug-sensitive[24].

In addition to *PIK3CA* mutations, EBV-positive tumours had frequent *ARID1A* (55%) and *BCOR* (23%) mutations and only rare *TP53* mutations. *BCOR*, encoding an anti-apoptotic protein, is also mutated in leukaemia[25] and medulloblastoma[26]. Among the CIN tumours, we observed *TP53* mutations in 71% of tumours. *CDH1* somatic mutations were enriched in the genomically stable subtype (37% of cases). *CDH1* germline mutations underlie hereditary diffuse gastric cancer (HDGC). However, germline analysis revealed only two *CDH1* polymorphisms, neither of which is known to be pathogenic. As in the EBV-subtype, inactivating *ARID1A* mutations were prevalent in the genomically stable subtype. We identified mutations of *RHOA* almost exclusively in genomically stable tumours, as discussed below.

We analysed the patterns of base changes within gastric cancer tumours and noted elevated rates of C to T transitions at CpG dinucleotides. We observed an elevated rate of A to C transversions at the 3′ adenine of AA dinucleotides, especially at AAG trinucleotides, as reported in oesophageal adenocarcinoma[27]. The A to C transversions were prominent in CIN, EBV and genomically stable, but as previously observed[27], not in MSI tumours (Supplementary Fig. 3.10).

We identified *RHOA* mutation in 16 cases, and these were enriched in the genomically stable subtype (15% of genomically stable cases, $P = 0.0039$). RHOA, when in the active GTP-bound form, acts through a variety of effectors, including ROCK1, mDIA and Protein Kinase N, to control actin-myosin-dependent cell contractility and cellular motility[28,29] and to activate STAT3 to promote tumorigenesis[30,31]. *RHOA* mutations were clustered in two adjacent amino-terminal regions that are predicted to be at the interface of RHOA with ROCK1 and other effectors (Fig. 4a, b). *RHOA* mutations were not at sites analogous to oncogenic mutations in RAS-family GTPases. Although one case harboured a codon 17 mutation, we did not identify the dominant-negative G17V mutations noted in T-cell neoplasms[32,33]. Rather, the mutations found in this study may act to modulate signalling downstream of RHOA. Biochemical studies found that the RHOA Y42C mutation attenuated activation of Protein Kinase N, without abrogated activation of mDia or ROCK1[34]. RHOA Y42, mutated in five tumours, corresponds to Y40 on HRAS, a residue which when mutated selectively reduces HRAS activation of RAF, but not other RAS effectors[35]. Given the role of RHOA in cell motility, modulation of RHOA may contribute to the disparate growth patterns and lack of cellular cohesion that are hallmarks of diffuse tumours.

Dysregulated RHO signalling was further implicated by the discovery of recurrent structural genomic alterations. Whole genome sequencing of 107 tumours revealed 5,696 structural rearrangements, including 74 predicted to produce in-frame gene fusions (Supplementary Information S3.7–3.8). *De novo* assembly of mRNA sequencing data confirmed 170 structural rearrangements (Supplementary Information S5.4a), including two cases with an interchromosomal translocation between *CLDN18* and *ARHGAP26* (*GRAF*). ARHGAP26 is a GTPase-activating protein (GAP) that facilitates conversion of RHO GTPases to the GDP state and has been implicated in enhancing cellular motility[34]. CLDN18

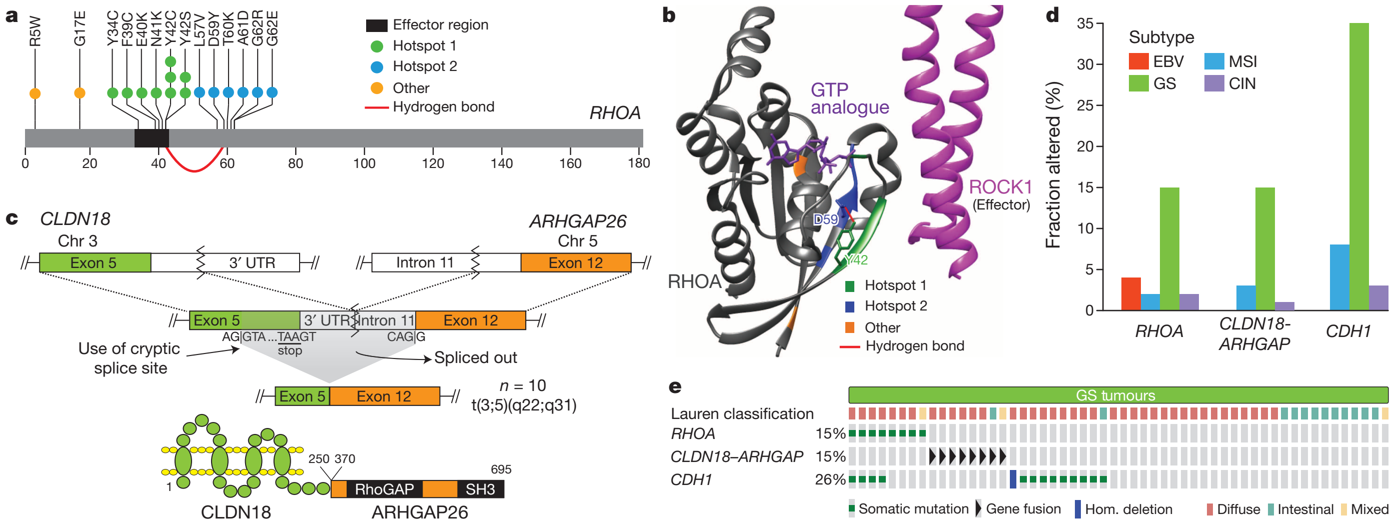


Figure 4 | ***RHOA* and *ARHGAP6/26* somatic genomic alterations are recurrent in genomically stable gastric cancer.** **a**, Missense mutations in the GTPase *RHOA*, including residues Y42 and D59, linked via hydrogen bond (red arc). **b**, Mutated regions (coloured as in panel **a**) mapped on the structures of RHOA and ROCK1. **c**, A schematic of *CLDN18–ARHGAP26* translocation is shown for the fusion transcript and predicted fusion protein. SH3 denotes SRC homology 3 domain. **d**, The frequency of *RHOA* and *CDH1* mutations, *CLDN18–ARHGAP6* or *ARHGAP26* fusions are shown across gastric cancer subtypes. **e**, *RHOA* mutations and *CLDN18–ARHGAP6* or *ARHGAP26* fusions are mutually exclusive in genomically stable tumours.

is a component of the tight junction adhesion structures[36]. RNA sequencing data from tumours without whole genome sequencing identified *CLDN18–ARHGAP26* fusions in 9 additional tumours, with two more cases showing *CLDN18* fusion to the homologous GAP encoded by *ARHGAP6* totalling 13 cases with these rearrangements (Supplementary Table 5.6).

The fusions linked exon 5 of *CLDN18* to exon 2 ($n = 2$) of *ARHGAP6*, to exon 10 ($n = 1$), or to exon 12 ($n = 10$) of *ARHGAP26* (Fig. 4c). As these fusions occur downstream of the *CLDN18* exon 5 stop codon, they appeared unlikely to enable translation of fusion proteins. However, mRNA sequencing revealed a mature fusion transcript in which the *ARHGAP26* or *ARHGAP6* splice acceptor activates a cryptic splice site within exon 5 of *CLDN18*, before the stop codon, yielding an in-frame fusion predicted to maintain the transmembrane domains of CLDN18 while fusing a large segment of ARHGAP26 or ARHGAP6 to the cytoplasmic carboxy terminus of CLDN18. These chimaeric proteins retain the carboxy-terminal GAP domain of ARHGAP26/6, potentially affecting ARHGAP's regulation of RHOA and/or cell motility. Furthermore, these fusions may also disrupt wild-type CLDN18, impacting cellular adhesion. The *CLDN18–ARHGAP* fusions were mutually exclusive with *RHOA* mutations and were enriched in genomically stable tumours (62%, $P = 10^{-3}$) (Fig. 4d). Within the genomically stable subtype, 30% of cases had either *RHOA* or *CLDN18–ARHGAP* alterations. Evaluation of gene expression status in pathways putatively regulated by RHOA using the Paradigm-Shift algorithm predicted activation of RHOA-driven pathways (Supplementary Fig. 11.4a–c), suggesting that these genomic aberrations contribute to the invasive phenotype of diffuse gastric cancer.

SCNA analysis using GISTIC identified 30 focal amplifications, 45 focal deletions, and chromosome arms subject to frequent alteration (Supplementary Figs 2.3–2.9). Focal amplifications targeted oncogenes such as *ERBB2, CCNE1, KRAS, MYC, EGFR, CDK6, GATA4, GATA6* and *ZNF217*. Additionally, we saw amplification of the gene that encodes the gastric stem cell marker CD44 and a novel recurrent amplification at 9p24.1 at the locus containing *JAK2, CD274* and *PDCD1LG2*. *JAK2* encodes a receptor tyrosine kinase and potential therapeutic target. *CD274* and *PDCD1LG2* encode PD-L1 and PD-L2, immunosuppressant proteins currently being evaluated as targets to augment anti-tumour immune response. Notably, these 9p amplifications were enriched in the EBV subgroup (15% of tumours), consistent with studies showing elevated PD-L1 expression in EBV-positive lymphoid cancers[37,38]. Evaluation of mRNA revealed elevated expression of *JAK2, PD-L1* and *PD-L2* in amplified cases (Supplementary Fig. 2.10). More broadly, PD-L1/2 expression was elevated in EBV-positive tumours, suggesting that PD-L1/2 antagonists and JAK2 inhibitors be tested in this subgroup. Focal deletions were identified at the loci of tumour suppressors such as *PTEN, SMAD4, CDKN2A* and *ARID1A*. Additional GISTIC analysis on the four molecular subtypes is detailed in Supplementary Figs 2.5–2.6.

Gene expression and proteomic analysis

Our analysis of each of the expression platforms revealed four mRNA, five miRNA and three RPPA clusters (Supplementary Methods S5–S7). Some expression clusters are similar across platforms (Supplementary Methods S10) and/or have correspondence with specific molecular subtypes. For example, mRNA cluster 1, miRNA cluster 4 and RPPA cluster 1 have substantial overlap and are strongly associated with genomically stable tumours, both individually and as a group; the 34 cases with all three assignments were predominantly genomically stable (20/34, $P = 2 \times 10^{-8}$). Similarly, mRNA cluster 3, miRNA cluster 2 and RPPA cluster 3 are similar and are associated with the MSI subtype as a group (12/22, $P = 5 \times 10^{-4}$). However, absolute correspondence between expression clusters and molecular subtypes was not always seen. For example, RPPA cluster 3 showed moderate association with both MSI and EBV ($P = 0.018$ and $P = 0.038$, respectively), and miRNA clusters each had similar proportions of CIN (no associations with $P < 0.05$). Overall, the expression data recapitulate features of the molecular classification, pointing to robustness of this taxonomy.

We analysed mRNA sequence data for alternative splicing events, finding *MET* exon 2 skipping in 82 of 272 (30%) cases, associated with increased MET expression ($P = 10^{-4}$). We also found novel variants of *MET* in which exons 18 and/or 19 were skipped (47/272; 17%; Supplementary Fig. 5.5). Intriguingly, the exons removed by these alterations encode regions of the kinase domain.

Through supervised analysis of RPPA data, we observed 45 proteins whose expression or phosphorylation was associated with the four molecular subtypes (Supplementary Fig. 7.2). Phosphorylation of EGFR (pY1068) was significantly elevated in the CIN subtype, consistent with amplification of *EGFR* within that subtype. We also found elevated expression of p53, consistent with frequent *TP53* mutation and aneuploidy in the CIN subtype.

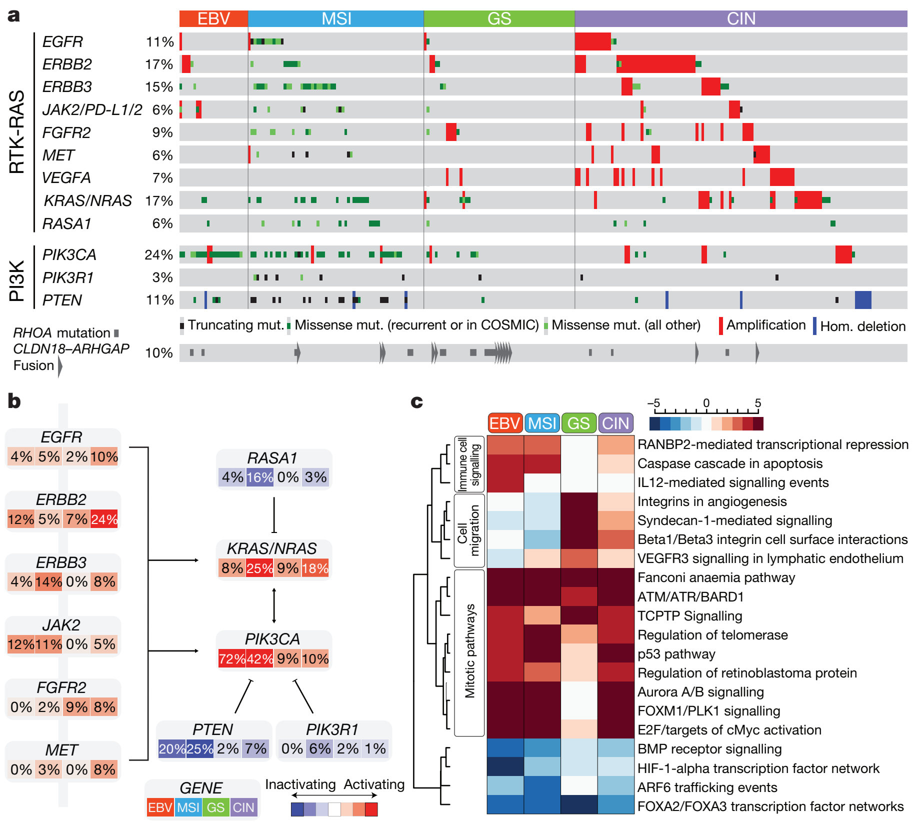


Figure 5 | Integrated molecular description of gastric cancer. **a**, Mutations, copy-number changes and translocations for select genes are shown across samples organized by molecular subtypes. Mutations that are recurrent in this data set or in the COSMIC repository are distinguished by colour. Alteration frequencies are expressed as a percentage of all cases. **b**, Alterations in RTK/RAS and RTK/PI(3)K signalling pathways across molecular subtypes. Red denotes predicted activation; blue denotes predicted inactivation. **c**, The heatmap shows NCI-PID pathways that are significantly elevated (red) or decreased (blue) in each of the four subtypes as compared with non-malignant gastric mucosa.

Integrated pathway analysis

We integrated SCNA and mutation data to characterize genomic alterations in known signalling pathways, including candidate therapeutic targets (Fig. 5a, b). We focused on alterations in receptor tyrosine kinases (RTKs) and RAS and PI(3)-kinase signalling. EBV-positive tumours contained *PIK3CA* mutations and recurrent *JAK2* and *ERBB2* amplifications. Although MSI cases generally lacked targetable amplifications, mutations in *PIK3CA*, *ERBB3*, *ERBB2* and *EGFR* were noted, with many mutations at 'hotspot' sites seen in other cancers (Supplementary Fig. 11.14). Absent from MSI gastric cancers were *BRAF* V600E mutations, commonly seen in MSI colorectal cancer[39]. Although the genomically stable subtype exhibited recurrent *RHOA* and *CLDN18* events, few other clear treatment targets were observed. In CIN tumours, we identified genomic amplifications of RTKs, many of which are amenable to blockade by therapeutics in current use or in development. Recurrent amplification of the gene encoding ligand VEGFA was notable given the gastric cancer activity of the VEGFR2 targeting antibody ramucirumab[40]. Additionally, frequent amplifications of cell cycle mediators (*CCNE1, CCND1* and *CDK6*) suggest the potential for therapeutic inhibition of cyclin-dependent kinases (Supplementary Fig. 11.15).

We compared expression within each subtype to that of the other subtypes, and to non-malignant gastric tissue ($n = 29$) (Supplementary Fig. 11.2). We computed an aggregate score for each pathway of the NCI pathway interaction database[41] and determined statistical significance by comparison with randomly generated pathways (Supplementary Methods S11). Hierarchical clustering of samples and pathways (Fig. 5c) revealed several notable patterns, including elevated expression of mitotic network components such as AURKA/B and E2F, targets of MYC activation, FOXM1 and PLK1 signalling and DNA damage response pathways across all subtypes, but to a lesser degree in genomically stable tumours. In contrast, the genomically stable subtype exhibited elevated expression of cell adhesion pathways, including the B1/B3 integrins, syndecan-1 mediated signalling, and angiogenesis-related pathways. These results suggest additional candidate therapeutic targets, including the aurora kinases (AURKA/B) and Polo-like (PLK) family members. The strength of IL-12 mediated signalling signatures in EBV-positive tumours suggests a robust immune cell presence. When coupled with evidence of PD-L1/2 overexpression, this finding adds rationale for testing immune checkpoint inhibitors in EBV-positive gastric cancer.

Discussion

Through this study of the molecular and genomic basis of gastric cancer, we describe a molecular classification (Fig. 6) that defines four major genomic subtypes of gastric cancer: EBV-infected tumours; MSI tumours; genomically stable tumours; and chromosomally unstable tumours. This

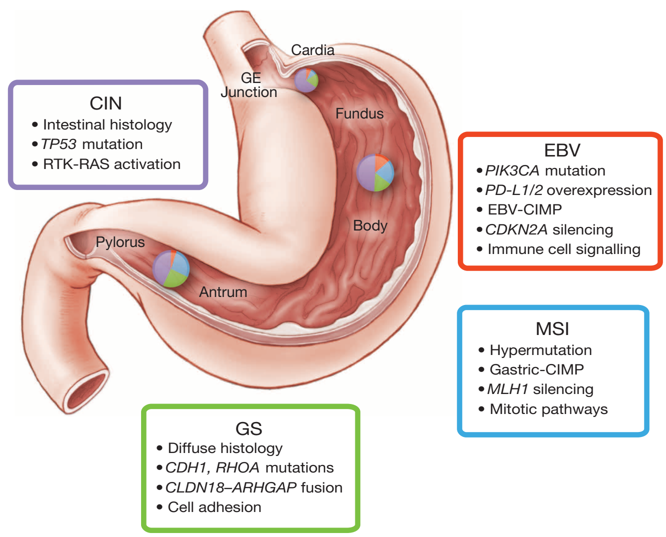


Figure 6 | Key features of gastric cancer subtypes. This schematic lists some of the salient features associated with each of the four molecular subtypes of gastric cancer. Distribution of molecular subtypes in tumours obtained from distinct regions of the stomach is represented by inset charts.

classification may serve as a valuable adjunct to histopathology. Importantly, these molecular subtypes showed distinct salient genomic features, providing a guide to targeted agents that should be evaluated in clinical trials for distinct populations of gastric cancer patients. Through existing testing for MSI and EBV and the use of emerging genomic assays that query focused gene sets for mutations and amplifications, the classification system developed through this study can be applied to new gastric cancer cases. We hope these results will facilitate the development of clinical trials to explore therapies in defined sets of patients, ultimately improving survival from this deadly disease.

METHODS SUMMARY

Fresh frozen gastric adenocarcinoma and matched germline DNA samples were obtained from 295 patients under IRB approved protocols. Genomic material and (when available) protein were subjected to single nucleotide polymorphism array somatic copy-number analysis, whole-exome sequencing, mRNA sequencing, miRNA sequencing, array-based DNA methylation profiling and reverse-phase protein arrays. A subset of samples was subjected to whole-genome sequencing. Initial analysis centred on the development of a classification scheme for gastric cancer. Subsequent analysis identified key features from each of the genomic/molecular platforms, looking both for features found across gastric cancer and those characteristic of individual gastric cancer subtypes. Primary and processed data are deposited at the Data Coordinating Center (https://tcga-data.nci.nih.gov/tcga/tcgaDownload.jsp); primary sequence files are deposited in CGHub (https://cghub.ucsc.edu/). Sample lists, and supporting data can be found at (https://tcga-data.nci.nih.gov/docs/publications/stad_2014/).

Received 21 February; accepted 13 May 2014.

Published online 23 July; corrected online 10 September 2014 (see full-text HTML version for details).

1. Ferlay, J. *et al.* GLOBOCAN 2012 v1.0, cancer incidence and mortality worldwide. IARC CancerBase **11**, http://globocan.iarc.fr, accessed on January 15, 2014 (2013).
2. Laurén, P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. *Acta Pathol., Microbiol. Scand.* **64,** 31–49 (1965).
3. *WHO Classification of Tumours of the Digestive System* 4th edn (International Agency for Research on Cancer, 2010).
4. Uemura, N. *et al. Helicobacter pylori* infection and the development of gastric cancer. *N. Engl. J. Med.* **345,** 784–789 (2001).
5. Bertuccio, P. *et al.* Recent patterns in gastric cancer: a global overview. *International J. Cancer* **125,** 666–673 (2009).
6. Richards, F. M. *et al.* Germline E-cadherin gene (*CDH1*) mutations predispose to familial gastric cancer and colorectal cancer. *Hum. Mol. Genet.* **8,** 607–610 (1999).
7. Keller, G. *et al.* Analysis for microsatellite instability and mutations of the DNA mismatch repair gene *hMLH1* in familial gastric cancer. *International J Cancer* **68,** 571–576 (1996).
8. Toyota, M. *et al.* Aberrant methylation in gastric cancer associated with the CpG island methylator phenotype. *Cancer Res.* **59,** 5438–5442 (1999).
9. Tan, I. B. *et al.* Intrinsic subtypes of gastric cancer, based on gene expression pattern, predict survival and respond differently to chemotherapy. *Gastroenterology* **141,** 476–485 (2011).
10. Lei, Z. *et al.* Identification of molecular subtypes of gastric cancer with different responses to PI3-kinase inhibitors and 5-fluorouracil. *Gastroenterology* **145,** 554–565 (2013).
11. Boussioutas, A. *et al.* Distinctive patterns of gene expression in premalignant gastric mucosa and gastric cancer. *Cancer Res.* **63,** 2569–2577 (2003).
12. Wang, K. *et al.* Exome sequencing identifies frequent mutation of *ARID1A* in molecular subtypes of gastric cancer. *Nature Genet.* **43,** 1219–1223 (2011).
13. Shen, R., Olshen, A. B. & Ladanyi, M. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. *Bioinformatics* **25,** 2906–2912 (2009).
14. Murphy, G., Pfeiffer, R., Camargo, M. C. & Rabkin, C. S. Meta-analysis shows that prevalence of Epstein-Barr virus-positive gastric cancer differs based on sex and anatomic location. *Gastroenterology* **137,** 824–833 (2009).
15. Matsusaka, K. *et al.* Classification of Epstein–Barr virus-positive gastric cancers by definition of DNA methylation epigenotypes. *Cancer Res.* **71,** 7187–7197 (2011).
16. Geddert, H., Zur Hausen, A., Gabbert, H. E. & Sarbia, M. EBV-infection in cardiac and non-cardiac gastric adenocarcinomas is associated with promoter methylation of p16, p14 and APC, but not hMLH1. *Anal. Cell. Pathol.* **33,** 143–149 (2010).
17. Lee, J. *et al.* High-throughput mutation profiling identifies frequent somatic mutations in advanced gastric adenocarcinoma. *PLoS ONE* **7,** e38892 (2012).
18. Sukawa, Y. *et al.* Alterations in the human epidermal growth factor receptor 2-phosphatidylinositol 3-kinase-v-Akt pathway in gastric cancer. *World J. Gastroenterology* **18,** 6577–6586 (2012).
19. Strong, M. J. *et al.* Differences in gastric carcinoma microenvironment stratify according to EBV infection intensity: implications for possible immune adjuvant therapy. *PLoS Pathog.* **9,** e1003341 (2013).
20. The Cancer Genome Atlas Network Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487,** 330–337 (2012).
21. Cancer Genome Atlas Research Network. Integrated genomic characterization of endometrial carcinoma. *Nature* **497,** 67–73 (2013).
22. Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499,** 214–218 (2013).
23. Bernal, M., Ruiz-Cabello, F., Concha, A., Paschen, A. & Garrido, F. Implication of the β2-microglobulin gene in the generation of tumor escape phenotypes. *Cancer Immunol. Immunother.* **61,** 1359–1371 (2012).
24. Greulich, H. *et al.* Functional analysis of receptor tyrosine kinase mutations in lung cancer identifies oncogenic extracellular domain mutations of *ERBB2*. *Proc. Natl Acad. Sci. USA* **109,** 14476–14481 (2012).
25. Grossmann, V. *et al.* Whole-exome sequencing identifies somatic mutations of *BCOR* in acute myeloid leukemia with normal karyotype. *Blood* **118,** 6153–6163 (2011).
26. Pugh, T. J. *et al.* Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature* **488,** 106–110 (2012).
27. Dulak, A. M. *et al.* Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity. *Nature Genet.* **45,** 478–486 (2013).
28. Ridley, A. J. *et al.* Cell migration: integrating signals from front to back. *Science* **302,** 1704–1709 (2003).
29. Thumkeo, D., Watanabe, S. & Narumiya, S. Physiological roles of Rho and Rho effectors in mammals. *Eur. J. Cell Biol.* **92,** 303–315 (2013).
30. Aznar, S. *et al.* Simultaneous tyrosine and serine phosphorylation of STAT3 transcription factor is involved in Rho A GTPase oncogenic transformation. *Mol. Biol. Cell* **12,** 3282–3294 (2001).
31. Yu, H. & Jove, R. The STATs of cancer–new molecular targets come of age. *Nature Rev. Cancer* **4,** 97–105 (2004).
32. Palomero, T. *et al.* Recurrent mutations in epigenetic regulators, *RHOA* and *FYN* kinase in peripheral T cell lymphomas. *Nature Genet.* **46,** 166–170 (2014).
33. Sakata-Yanagimoto, M. *et al.* Somatic *RHOA* mutation in angioimmunoblastic T cell lymphoma. *Nature Genet.* **46,** 171–175 (2014).
34. Doherty, G. J. *et al.* The endocytic protein GRAF1 is directed to cell-matrix adhesion sites and regulates cell spreading. *Mol. Biol. Cell* **22,** 4380–4389 (2011).
35. Joneson, T., White, M. A., Wigler, M. H. & Bar-Sagi, D. Stimulation of membrane ruffling and MAP kinase activation by distinct effectors of RAS. *Science* **271,** 810–812 (1996).
36. Türeci, O. *et al.* Claudin-18 gene structure, regulation, and expression is evolutionary conserved in mammals. *Gene* **481,** 83–92 (2011).
37. Chen, B. J. *et al.* PD-L1 expression is characteristic of a subset of aggressive B-cell lymphomas and virus-associated malignancies. *Clinical Cancer Res* **19,** 3462–3473 (2013).
38. Green, M. R. *et al.* Constitutive AP-1 activity and EBV infection induce PD-L1 in Hodgkin lymphomas and posttransplant lymphoproliferative disorders: implications for targeted therapy. *Clinical Cancer Res* **18,** 1611–1618 (2012).
39. The Cancer Genome Atlas Network Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487,** 330–337 (2012).
40. Fuchs, C. S. *et al.* Ramucirumab monotherapy for previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (REGARD): an international, randomised, multicentre, placebo-controlled, phase 3 trial. *Lancet* **383,** 31–39 (2014).
41. Schaefer, C. F. *et al.* PID: the pathway interaction database. *Nucleic Acids Res.* **37,** D674–D679 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to all the patients and families who contributed to this study and to C. Gunter and J. Weinstein for scientific editing, to M. Sheth for administrative support and to L. Omberg for support with the Sage Bionetworks Synapse platform. This work was supported by the Intramural Research Program and the following grants from the United States National Institutes of Health: 5U24CA143799, 5U24CA143835, 5U24CA143840, 5U24CA143843, 5U24CA143845, 5U24CA143848, 5U24CA143858, 5U24CA143866, 5U24CA143867, 5U24CA143882, 5U24CA143883, 5U24CA144025, U54HG003067, U54HG003079, U54HG003273 and P30CA16672.

Author Contributions The Cancer Genome Atlas Research Network contributed collectively to this study. Biospecimens were provided by the tissue source sites and processed by the Biospecimen Core Resource. Data generation and analyses were performed by the genome-sequencing centres, cancer genome-characterization centres and genome data analysis centres. All data were released through the Data Coordinating Center. The NCI and NHGRI project teams coordinated project activities. The following TCGA investigators of the Stomach Analysis Working Group contributed substantially to the analysis and writing of this manuscript. Project leaders, A. J. Bass, P. W. Laird, I. Shmulevich; data coordinator, V. Thorsson; analysis coordinators, V. Thorsson, N. Schultz; manuscript coordinator, M. Sheth; graphics coordinator, T. Hinoue; DNA sequence analysis, A. Taylor-Weiner, A. Pantazi, C. Sougnez, K. Kasaian; mRNA analysis, R. Bowlby, A. J. Mungall; miRNA analysis, A. Chu, A. Gordon Robertson, D. Yang; DNA methylation analysis, T. Hinoue, H. Shen, P. W. Laird; copy number analysis, A. Cherniack; protein analysis, J.-S. Lee, R. Akbani; pathway/integrated analysis, N. Weinhold, S. Reynolds, C. Curtis, R. Shen, S. Ng, B. Raphael, H.-T. Wu, Y. Liu, V. Thorsson, N. Schultz; pathology expertise and clinical data, A. Boussioutas, B. G. Schneider, J. Kim, J. E. Willis, M. L. Gulley, K. Garman, M. Blanca Piazuelo, V. Thorsson, K. M. Leraas, T. Lichtenberg, J. A. Demchok, A. J. Bass; microbiome analysis, C. S. Rabkin, M. L. Gulley, R. Bowlby, A. J. Mungall, A. Chu and C. Pedamallu.

Author Information The primary and processed data used to generate the analyses presented here can be downloaded from The Cancer Genome Atlas at (https://tcga-data.nci.nih.gov/tcga/tcgaDownload.jsp). All of the primary sequence files are deposited in CGHub and all other data are deposited at the Data Coordinating Center (DCC) for public access (http://cancergenome.nih.gov/) and (https://cghub.ucsc.edu/). Additional sample data and supporting data are available from (https://tcga-data.nci.nih.gov/docs/publications/stad_2014/). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.J.B. (adam_bass@dfci.harvard.edu).

The Cancer Genome Atlas Research Network

Analysis Working Group: Dana-Farber Cancer Institute Adam J. Bass[1]; **Institute for Systems Biology** Vesteinn Thorsson[2], Ilya Shmulevich[2], Sheila M. Reynolds[2], Michael Miller[2], Brady Bernard[2]; **University of Southern California** Toshinori Hinoue[3], Peter W. Laird[3], Christina Curtis[4], Hui Shen[3], Daniel J. Weisenberger[3]; **Memorial Sloan Kettering Cancer Center** Nikolaus Schultz[5], Ronglai Shen[6], Nils Weinhold[5], David P. Kelsen[7]; **BC Cancer Agency** Reanne Bowlby[8], Andy Chu[8], Katayoon Kasaian[8], Andrew J. Mungall[8], A. Gordon Robertson[8], Payal Sipahimalani[8]; **The Eli & Edythe L. Broad Institute** Andrew D. Cherniack[9], Gad Getz[9], Yingchun Liu[9], Michael S. Noble[9], Chandra Pedamallu[9], Carrie Sougnez[9], Amaro Taylor-Weiner[9]; **MD Anderson Cancer Center** Rehan Akbani[10], Ju-Seog Lee[10], Wenbin Liu[10], Gordon B. Mills[11], Da Yang[12], Wei Zhang[12]; **Harvard Medical School** Angeliki Pantazi[13], Michael Parfenov[13]; **University of North Carolina** Margaret Gulley[14]; **Vanderbilt University** M. Blanca Piazuelo[15], Barbara G. Schneider[15]; **Asan Medical Center** Jihun Kim[16]; **University of Melbourne** Alex Boussioutas[17]; **National Cancer Institute** Margi Sheth[18], John A. Demchok[18], Charles S. Rabkin[19]; **Case Western Reserve University** Joseph E. Willis[20]; **University of California at Santa Cruz** Sam Ng[21]; **Duke University** Katherine Garman[22]; **University of Michigan** David G. Beer[23]; **University of Pittsburgh** Arjun Pennathur[24]; **Brown University** Benjamin J. Raphael[25], Hsin-Ta Wu[25]; **Brigham and Women's Hospital** Robert Odze[26]; **National Cancer Center** Hark K. Kim[27]; **Nationwide Children's Hospital** Jay Bowen[28], Kristen M. Leraas[28], Tara M. Lichtenberg[28], Stephanie Weaver[28]; **Washington University** Michael McLellan[29]; **Greater Poland Cancer Centre** Maciej Wiznerowicz[30], **KU Leuven:** Ryo Sakai[31]

Genome Sequencing Center: The Eli & Edythe L. Broad Institute Gad Getz[9], Carrie Sougnez[9], Michael S. Lawrence[9], Kristian Cibulskis[9], Lee Lichtenstein[9], Sheila Fisher[9], Stacey B. Gabriel[9], Eric S. Lander[9]; **Washington University in St. Louis** Li Ding[29], Beifang Niu[29]

Genome Characterization Centers: BC Cancer Agency Adrian Ally[8], Miruna Balasundaram[8], Inanc Birol[8], Reanne Bowlby[8], Denise Brooks[8], Yaron S. N. Butterfield[8], Rebecca Carlsen[8], Andy Chu[8], Justin Chu[8], Eric Chuah[8], Hye-Jung E. Chun[8], Amanda Clarke[8], Noreen Dhalla[8], Ranabir Guin[8], Robert A. Holt[8], Steven J. M. Jones[8], Katayoon Kasaian[8], Darlene Lee[8], Haiyan A. Li[8], Emilia Lim[8], Yussanne Ma[8], Marco A. Marra[8], Michael Mayo[8], Richard A. Moore[8], Andrew J. Mungall[8], Karen L. Mungall[8], Ka Ming Nip[8], A. Gordon Robertson[8], Jacqueline E. Schein[8], Payal Sipahimalani[8], Angela Tam[8], Nina Thiessen[8]; **The Eli & Edythe L. Broad Institute** Rameen Beroukhim[9], Scott L. Carter[9], Andrew D. Cherniack[9], Juok Cho[9], Kristian Cibulskis[9], Daniel DiCara[9], Scott Frazer[9], Sheila Fisher[9], Stacey B. Gabriel[9], Nils Gehlenborg[9], David I. Heiman[9], Joonil Jung[9], Jaegil Kim[9], Eric S. Lander[9], Michael S. Lawrence[9], Lee Lichtenstein[9], Pei Lin[9], Matthew Meyerson[9], Akinyemi I. Ojesina[9], Chandra Sekhar Pedamallu[9], Gordon Saksena[9], Steven E. Schumacher [9], Carrie Sougnez[9], Petar Stojanov[9], Barbara Tabak[9], Amaro Taylor-Weiner[9], Doug Voet[9], Mara Rosenberg[9], Travis I. Zack[9], Hailei Zhang[9], Lihua Zou[9]; **Harvard Medical School/Brigham & Women's Hospital/MD Anderson Cancer Center** Alexei Protopopov[32], Netty Santoso[13], Michael Parfenov[13], Semin Lee[33], Jianhua Zhang[32], Harshad S. Mahadeshwar[32], Jiabin Tang[32], Xiaojia Ren[13], Sahil Seth[32], Lixing Yang[33], Andrew W. Xu[33], Xingzhi Song[32], Angeliki Pantazi[13], Ruibin Xi[33], Christopher A. Bristow[32], Angela Hadjipanayis[13], Jonathan Seidman[13], Lynda Chin[32], Peter J. Park[33], Raju Kucherlapati[13]; **MD Anderson Cancer Center** Rehan Akbani[10], Shiyun Ling[10], Wenbin Liu[10], Arvind Rao[10], John N. Weinstein[10], Sang-Bae Kim[11], Ju-Seog Lee[11], Yiling Lu[11], Gordon Mills[11]; **University of Southern California Epigenome Center** Peter W. Laird[3], Toshinori Hinoue[3], Daniel J. Weisenberger[3], Moiz S. Bootwalla[3], Phillip H. Lai[3], Hui Shen[3], Timothy Triche Jr[3], David J. Van Den Berg[3]; **The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University** Stephen B. Baylin[34], James G. Herman[34]

Genome Data Analysis Centers: The Eli & Edythe L. Broad Institute Gad Getz[9], Lynda Chin[32], Yingchun Liu[9], Bradley A. Murray[9], Michael S. Noble[9]; **Memorial Sloan-Kettering Cancer Center** B. Arman Askoy[5], Giovanni Ciriello[5], Gideon Dresdner[5], Jianjiong Gao[5], Benjamin Gross[5], Anders Jacobsen[5], William Lee[5], Ricardo Ramirez[5], Chris Sander[5], Nikolaus Schultz[5], Yasin Senbabaoglu[5], Rileen Sinha[5], S. Onur Sumer[5], Yichao Sun[5], Nils Weinhold[5]; **Institute for Systems Biology** Vésteinn Thorsson[2], Brady Bernard[2], Lisa Iype[2], Roger W. Kramer[2], Richard Kreisberg[2], Michael Miller[2], Sheila M. Reynolds[2], Hector Rovira[2], Natalie Tasman[2], Ilya Shmulevich[2]; **University of California, Santa Cruz** Sam Ng[21], David Haussler[21], Josh M. Stuart[21]; **MD Anderson Cancer Center** Rehan Akbani[10], Shiyun Ling[10], Wenbin Liu[10], Arvind Rao[10], John N. Weinstein[10], Roeland G. W. Verhaak[32], Gordon B. Mills[11]; **Brown University** Mark D. M. Leiserson[25], Benjamin J. Raphael[25], Hsin-Ta Wu[25]; **University of California San Francisco** Barry S. Taylor[35]

Biospecimen Core Resource: The Research Institute at Nationwide Children's Hospital Aaron D. Black[28], Jay Bowen[28], Julie Ann Carney[28], Julie M. Gastier-Foster[28], Carmen Helsel[28], Kristen M. Leraas[28], Tara M. Lichtenberg[28], Cynthia McAllister[28], Nilsa C. Ramirez[28], Teresa R. Tabler[28], Lisa Wise[28], Erik Zmuda[28]; **International Genomics Consortium** Robert Penny[36], Daniel Crain[36], Johanna Gardner[36], Kevin Lau[36], Erin Curely[36], David Mallery[36], Scott Morris[36], Joseph Paulauskis[36], Troy Shelton[36], Candace Shelton[36], Mark Sherman[36]

Tissue Source Sites: Buck Institute for Research on Aging Christopher Benz[37]; **Chonnam National University Medical School** Jae-Hyuk Lee[38]; **City Clinical Oncology Dispensary** Konstantin Fedosenko [39], Georgy Manikhas[39]; **Cureline** Olga Potapova[40], Olga Voronina[40], Dmitry Belyaev[40], Oleg Dolzhansky[40]; **UNC Lineberger Comprehensive Cancer Center** W. Kimryn Rathmell[41]; **Greater Poland Cancer Centre** Jakub Brzezinski[30], Matthew Ibbs[30], Konstanty Korski[30], Witold Kycler[30], Radoslaw Łaźniak[30], Ewa Leporowska[30], Andrzej Mackiewicz[30], Dawid Murawa[30], Pawel Murawa[30], Arkadiusz Spychała[30], Wiktoria M. Suchorska[30], Honorata Tatka[30], Marek Teresiak[30], Maciej Wiznerowicz[30]; **Helen F. Graham Cancer Center & Research Institute** Raafat Abdel-Misih[42], Joseph Bennett[42], Jennifer Brown[42], Mary Iacocca[42], Brenda Rabeno[42]; **Keimyung University School of Medicine** Sun-Young Kwon[43]; **International Genomics Consortium** Robert Penny[36], Johanna Gardner[36], Ariane Kemkes[36], David Mallery[36], Scott Morris[36], Troy Shelton[36], Candace Shelton[36], Erin Curley[36]; **Ontario Tumour Bank** Iakovina Alexopoulou[44], Jay Engel[45], John Bartlett[46], Monique Albert[46]; **Pusan National University Hospital** Do-Youn Park[47]; **University of Pittsburgh School of Medicine** Rajiv Dhir[24], James Luketich[24], Rodney Landreneau[24]

Disease Working Group: Memorial Sloan-Kettering Cancer Center Yelena Y. Janjigian[7], David P. Kelsen[7], Eunjung Cho[48], Marc Ladanyi[48], Laura Tang[48]; **Duke University** Shannon J. McCall[49]; **Asan Medical Center** Young S. Park[16]; **Yonsei University College of Medicine** Jae-Ho Cheong[50]; **MD Anderson Cancer Center** Jaffer Ajani[51]; **National Cancer Institute** M. Constanza Camargo[19]

Data Coordination Center: SRA International Shelley Alonso[52], Brenda Ayala[52], Mark A. Jensen[52], Todd Pihl[52], Rohini Raman[52], Jessica Walton[52], Yunhu Wan[52]

Project Team: National Cancer Institute John A. Demchok[18], Greg Eley[18], Kenna R. Mills Shaw[18], Margi Sheth[18], Roy Tarnuzzer[18], Zhining Wang[18], Liming Yang[18], Jean Claude Zenklusen[18], Tanja Davidsen[53], Carolyn M. Hutter[54], Heidi J. Sofia[54]; **SAIC-Frederick** Robert Burton[55], Sudha Chudamani[55] & Jia Liu[55]

[1]Department of Medical Oncology and the Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA. [2]Institute for Systems Biology, Seattle, Washington 98109, USA. [3]USC Epigenome Center, University of Southern California, Los Angeles, California 90033, USA. [4]University of Southern California, Department of Preventive Medicine, USC/Norris Comprehensive Cancer Center, Los Angeles, California 90033, USA. [5]Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [6]Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [7]Department of Medicine, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA. [8]Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, BC V5Z 4S6, Canada. [9]The Eli and Edythe L. Broad Institute, Cambridge, Massachusetts 02142, USA. [10]Department of Bioinformatics and Computational Biology, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [11]Department of Systems Biology, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [12]Department of Pathology, University of Texas MD Anderson Cancer Center, Texas 77030, USA. [13]Department of Medicine, Harvard Medical School, Boston, Massachusetts 02215, USA. [14]Department of Pathology and Laboratory Medicine, University of North Carolina-Chapel Hill, Chapel Hill, North Carolina 27599, USA. [15]Department of Medicine, Vanderbilt University Medical Center, 2215 Garland Avenue, Nashville, Tennessee 37232, USA. [16]Department of Pathology, University of Ulsan College of Medicine, Asan Medical Center, Seoul, 138-736, South Korea. [17]Sir Peter MacCallum Cancer Department of Oncology, University of Melbourne, East Melbourne 3002, Australia. [18]National Cancer Institute, Bethesda, Maryland 20892, USA. [19]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland 20892, USA. [20]Department of Pathology, Case Western Reserve University, Cleveland, Ohio 44106, USA. [21]Department of Biomolecular Engineering and Center for Biomolecular Science and Engineering, University of California-Santa Cruz, Santa Cruz, California 95064, USA. [22]Division of Gastroenterology, Department of Medicine, Duke University, Durham, North Carolina 27710, USA. [23]Department of Thoracic Surgery, University of Michigan Cancer Center, Ann Arbor, Michigan 48109, USA. [24]University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA. [25]Department of Computer Science & Center for Computational Molecular Biology, Brown University, 115 Waterman Street, Providence, Rhode Island 02912, USA. [26]Department of Pathology, Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115, USA. [27]National Cancer Center, Goyang, 410-769, Republic of

Korea. [28]The Research Institute at Nationwide Children's Hospital, Columbus, Ohio 43205, USA. [29]The Genome Institute, Washington University, St Louis, Missouri 63108, USA. [30]Greater Poland Cancer Centre, Garbary, 15, 61-866, Poznan, Poland. [31]KU Leuven, Department of Electrical Engineering-ESAT (STADIUS), Leuven, Belgium. [32]Institute for Applied Cancer Science, Department of Genomic Medicine, University of Texas MD Anderson Cancer Center, Houston, Texas 77054, USA. [33]The Center for Biomedical Informatics, Harvard Medical School, Boston, Massachusetts 02115, USA. [34]Cancer Biology Division, Johns Hopkins University, Baltimore, Maryland 21231, USA. [35]Helen Diller Family Comprehensive Cancer Center, University of California-San Francisco, San Francisco, California 94143-0128, USA. [36]International Genomics Consortium, Phoenix, Arizona 85004, USA. [37]Buck Institute for Research on Aging, Novato, California 94945, USA. [38]Chonnam National University Medical School, Gwangju, 501-746, Republic of Korea. [39]City Clinical Oncology Dispensary, Saint Petersburg 198255, Russia. [40]Cureline, Inc., South San Francisco, California 94080, USA. [41]Departments of Medicine and Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [42]Helen F. Graham Cancer Center & Research Institute, Christiana Care Health System, Newark, Delaware 19713, USA. [43]Keimyung University School of Medicine, Daegu, 700-712, Republic of Korea. [44]Ontario Tumour Bank - Hamilton site, St. Joseph's Healthcare Hamilton, Hamilton, Ontario L8N 3Z5, Canada. [45]Ontario Tumour Bank - Kingston site, Kingston General Hospital, Kingston, Ontario K7L 5H6, Canada. [46]Ontario Tumour Bank, Ontario Institute for Cancer Research, Toronto, Ontario M5G 0A3, Canada. [47]Pusan National University Hospital, Busan, 602-739, Republic of Korea. [48]Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [49]Department of Pathology, Duke University, Durham, North Carolina 27710, USA. [50]Department of Surgery, Yonsei University College of Medicine, Seoul, 120-752, Republic of Korea. [51]Department of Gastrointestinal Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [52]SRA International, Fairfax, Virginia 22033, USA. [53]Center for Biomedical Informatics and Information Technology, National Cancer Institute, Rockville, Maryland 20850, USA. [54]National Human Genome Research Institute, Bethesda, Maryland 20892, USA. [55]SAIC-Frederick, Inc., Frederick, Maryland 21702, USA.

LETTER

doi:10.1038/nature13712

The diversity of quasars unified by accretion and orientation

Yue Shen[1,2] & Luis C. Ho[2,3]

Quasars are rapidly accreting supermassive black holes at the centres of massive galaxies. They display a broad range of properties across all wavelengths, reflecting the diversity in the physical conditions of the regions close to the central engine. These properties, however, are not random, but form well-defined trends. The dominant trend is known as 'Eigenvector 1', in which many properties correlate with the strength of optical iron and [O III] emission[1–3]. The main physical driver of Eigenvector 1 has long been suspected[4] to be the quasar luminosity normalized by the mass of the hole (the 'Eddington ratio'), which is an important parameter of the black hole accretion process. But a definitive proof has been missing. Here we report an analysis of archival data that reveals that the Eddington ratio indeed drives Eigenvector 1. We also find that orientation plays a significant role in determining the observed kinematics of the gas in the broad-line region, implying a flattened, disk-like geometry for the fast-moving clouds close to the black hole. Our results show that most of the diversity of quasar phenomenology can be unified using two simple quantities: Eddington ratio and orientation.

The optical and ultraviolet spectra of quasars show emission lines with a wide variety of strengths (equivalent widths) and velocity widths. However, despite their great diversity in outward appearance, quasars possess surprising regularity in their physical properties. A seminal principal-component analysis[1] of 87 low-redshift broad-line quasars discovered that the main variance (Eigenvector 1, or EV1) in their optical properties arises from an anti-correlation between the strength of the narrow [O III] $\lambda = 5{,}007$ Å and broad Fe II emission. Along with other properties that also correlate with Fe II strength[2,3,5], these observations establish EV1 as a physical sequence of broad-line quasar properties. In the two-dimensional plane of Fe II strength (measured by the ratio of Fe II equivalent width within 4,434–4,684 Å to broad Hβ equivalent width, $R_{\rm Fe\,II} \equiv {\rm EW}_{\rm Fe\,II}/{\rm EW}_{\rm H\beta}$) and the full-width at half-maximum of broad Hβ (${\rm FWHM}_{\rm H\beta}$), EV1 is defined as the horizontal trend with $R_{\rm Fe\,II}$, where the average [O III] strength and ${\rm FWHM}_{\rm H\beta}$ decrease[1,2]. Figure 1 shows the EV1 sequence for about 20,000 broad-line quasars drawn from the Sloan Digital Sky Survey (SDSS)[6,7] (see Supplementary Information for details of the sample).

The statistics of the SDSS quasar sample allows us to divide the sample into bins of $R_{\rm Fe\,II}$ and ${\rm FWHM}_{\rm H\beta}$ (the grey grid in Fig. 1) and study the average [O III] properties in each bin. Figure 2 shows the average [O III] line profiles in each bin, as a function of the quasar continuum luminosity L measured at 5,100 Å. In addition to the EV1 sequence, the [O III] strength also decreases with $L_{5{,}100\,\text{Å}}$, following the Baldwin effect[8,9] initially discovered for the broad C IV line[10]. The [O III] profile can be decomposed into a core component, centred consistently at the systemic redshift, and a blueshifted wing component. The core component strongly follows the EV1 and Baldwin trends, while the wing component only shows a mild decrement with L and $R_{\rm Fe\,II}$ (Supplementary Information and Extended Data Figs 1–2). This may suggest that the core component is mostly powered by photoionization from the quasar, while the wing component is excited by other mechanisms, such as shocks associated with outflows[11].

In addition to the strongest narrow [O III] lines, all other optical narrow forbidden lines (such as [Ne V], [Ne III], [O II] and [S II]) show similar EV1 trends and the Baldwin effect. Hot dust emission detected using WISE[12], presumably coming from a dusty torus[13,14], also increases with $R_{\rm Fe\,II}$. In the Supplementary Information (and Extended Data Figs 3–7) we summarize all updated and new observations that firmly establish the EV1 sequence.

The [O III]-emitting region is photoionized by the ionizing continuum from the accreting black hole. But the EV1 correlation of [O III] strength with $R_{\rm Fe\,II}$ holds even when optical luminosity is fixed, as demonstrated in Fig. 2. This suggests that another physical property of black hole accretion changes with $R_{\rm Fe\,II}$, one that, in turn, affects the relative contribution in the ionizing part of the quasar continuum as seen by the narrow-line region. The most likely possibility is the black hole mass $M_{\rm BH}$, or equivalently, the Eddington ratio $L/M_{\rm BH}$, given that L is fixed. The much less likely alternative would be that the [O III] narrow-line region changes as a function of $R_{\rm Fe\,II}$. Reverberation mapping studies of nearby active galactic nuclei (AGN)[15] have suggested that a virial estimate of $M_{\rm BH}$ may be derived by combining the broad-line region size $R_{\rm BLR}$ (measured from

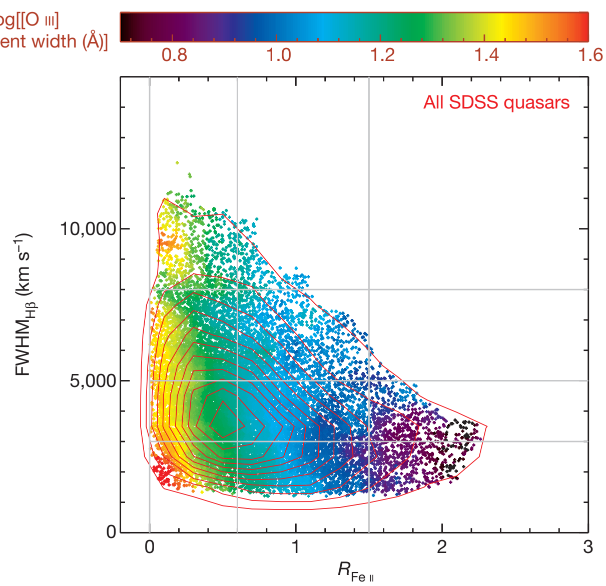


Figure 1 | Distribution of quasars in the EV1 plane. The horizontal axis is the relative Fe II strength, $R_{\rm Fe\,II}$, and the vertical axis is the broad Hβ FWHM. The red contours show the distribution of our SDSS quasar sample (with quasar density increasing from outer to inner contours), and the points show individual objects. We colour-code the points by the [O III] $\lambda = 5{,}007$ Å equivalent width, averaged over all nearby objects in a smoothing box of $\Delta R_{\rm Fe\,II} = 0.2$ and $\Delta{\rm FWHM}_{\rm H\beta} = 1{,}000\ {\rm km\,s^{-1}}$. The EV1 sequence[1] is the systematic trend of decreasing [O III] strength with increasing $R_{\rm Fe\,II}$. The grey grid divides this plane into bins of ${\rm FWHM}_{\rm H\beta}$ and $R_{\rm Fe\,II}$, in which we study the stacked spectral properties.

[1]Carnegie Observatories, 813 Santa Barbara Street, Pasadena, California 91101, USA. [2]Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China. [3]Department of Astronomy, School of Physics, Peking University, Beijing 100871, China.

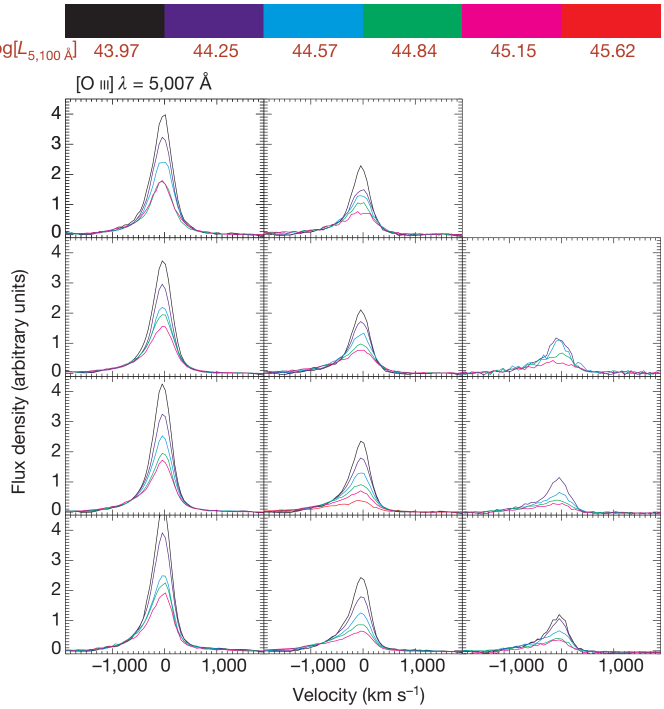


Figure 2 | Average [O III] profiles in the EV1 plane. Each panel shows the stacked [O III] $\lambda = 5{,}007$ Å line of quasars in the $R_{\mathrm{Fe\,II}}$–FWHM$_{\mathrm{H}\beta}$ bins defined by the grey grid in Fig. 1 (in the same layout). $R_{\mathrm{Fe\,II}}$ increases from left to right, and FWHM$_{\mathrm{H}\beta}$ increases from bottom to top. In each bin we further divided the quasars into different luminosity bins using the measured $L_{5,100\,Å}$ continuum luminosities. We have normalized the line fluxes by the (host-corrected) average quasar continuum luminosity $L_{5,100\,Å}$ for each stacking subset; hence, these stacked lines reflect the relative [O III] strength among different samples. In addition to the decrease of [O III] strength when $R_{\mathrm{Fe\,II}}$ increases (that is, Fig. 1), we also observe a decrease in [O III] strength with increasing luminosity[8,9]. The [O III] profile is in general asymmetric, with a blueshifted wing, whose relative contribution to the total profile increases when $R_{\mathrm{Fe\,II}}$ or luminosity increases.

the time lag between continuum and emission-line variability) and the virial velocity of the line-emitting clouds estimated from the linewidth: $M_{\mathrm{BH,virial}} \propto R_{\mathrm{BLR}}\mathrm{FWHM}^2_{\mathrm{H}\beta}/G$, where G is the gravitational constant. The average FWHM$_{\mathrm{H}\beta}$ does decrease by about 0.2 dex when $R_{\mathrm{Fe\,II}}$ increases from 0 to 2, and this fact underlies the earlier suggestion that EV1 is driven by the Eddington ratio[4,16].

A remarkable feature in Figs 1 and 2 is that the sequence is predominantly horizontal: there is little trend with FWHM$_{\mathrm{H}\beta}$ at fixed $R_{\mathrm{Fe\,II}}$. The standard virial mass estimators[15,17] would suggest that there is a strong vertical segregation in M_{BH}, by a factor of a few in the vertical bins in Fig. 1. If lower M_{BH} (or higher Eddington ratio) leads to weaker [O III], as in the EV1 relation (that is, the horizontal trend), we should also see a vertical trend in Fig. 1. The absence of such a trend suggests that there is substantial scatter between FWHM$_{\mathrm{H}\beta}$ and the actual virial velocity, and the vertical spread in FWHM$_{\mathrm{H}\beta}$ in the EV1 plane largely does not track the spread in true black hole masses.

We propose, instead, that the sequence in $R_{\mathrm{Fe\,II}}$ is driven by M_{BH}; but the dispersion in FWHM$_{\mathrm{H}\beta}$ at fixed $R_{\mathrm{Fe\,II}}$ is largely due to an orientation effect, as expected in a flattened broad-line region geometry. We first demonstrate that the average M_{BH} indeed decreases with $R_{\mathrm{Fe\,II}}$ for our quasar sample. We achieve this by measuring the clustering of SDSS quasars with low and high $R_{\mathrm{Fe\,II}}$ values. In the hierarchical clustering Universe, more massive galaxies (which contain more massive black holes) form in rarer density peaks and are more strongly clustered[18]. We therefore expect quasars with larger $R_{\mathrm{Fe\,II}}$ are less strongly clustered. This exercise, however, requires a very large sample size to achieve statistically significant results and has not been possible until now. Here we take advantage of the largest spectroscopic sample of galaxies from SDSS-III[19], and use the much larger (by a factor of about 40) galaxy sample to cross-correlate[20] with our quasar sample at redshift $z \approx 0.5$ to substantially improve the clustering measurements. The resulting cross-correlation functions are shown in Fig. 3a, for the two quasar subsamples divided at the median $R_{\mathrm{Fe\,II}}$. A significant clustering difference is detected at 3.48σ: quasars with larger $R_{\mathrm{Fe\,II}}$ are indeed less strongly clustered, confirming that they have on average lower M_{BH}.

In the EV1 plane (Fig. 1), the distribution in FWHM$_{\mathrm{H}\beta}$ at fixed $R_{\mathrm{Fe\,II}}$ is roughly log-normal, with mean value decreasing with $R_{\mathrm{Fe\,II}}$ and a dispersion of about 0.2 dex (Extended Data Fig. 8). We argued above that this dispersion is largely due to orientation-induced FWHM variations in the case of a flattened broad-line region geometry. For a small subset of quasars that are radio-loud (around 10% of the population), it is possible to infer the orientation of the accretion disk, and by extension, the broad-line region, using resolved radio morphology to deduce the orientation of the jet. Such studies[21,22] show that high-inclination (more edge-on)

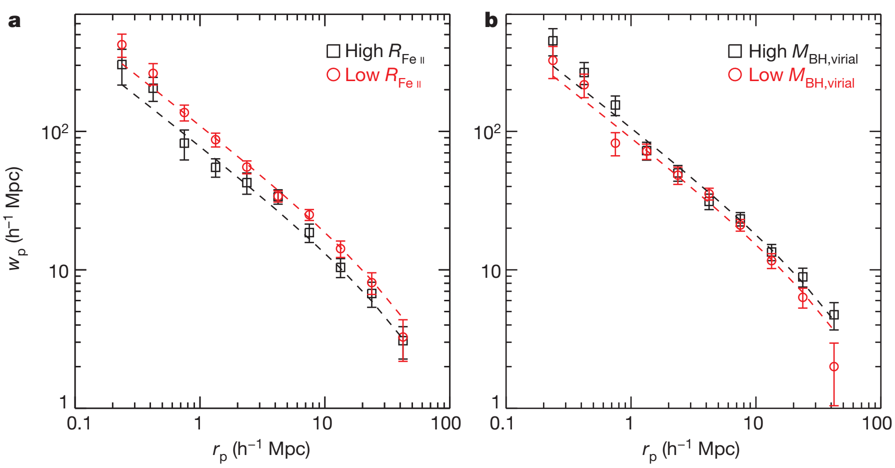


Figure 3 | Cross-correlation functions between different quasar subsamples and a galaxy sample. r_p is the transverse comoving separation and w_p is the projected two-point correlation function. **a**, Difference in the clustering strength when the quasar sample is divided at the median $R_{\mathrm{Fe\,II}}$. A significant difference (3.48σ) is detected: quasars with stronger $R_{\mathrm{Fe\,II}}$ are less strongly clustered, indicating they have on average smaller black hole masses. **b**, Difference in the clustering strength when the quasar sample is divided by the virial black hole mass estimates based on FWHM$_{\mathrm{H}\beta}$. No significant difference (1.64σ) is detected, indicating there is substantial overlap in the actual black hole masses between the two subsamples owing to the uncertainties in these FWHM-based virial black hole masses. Orientation-induced FWHM$_{\mathrm{H}\beta}$ dispersion can naturally lead to such uncertainties. Error bars are 1σ measurement errors estimated with jackknife resampling (Supplementary Information).

broad-line radio quasars have on average larger $\mathrm{FWHM}_{\mathrm{H}\beta}$, in accordance with the orientation hypothesis. Below, we perform a different test for the more general radio-quiet quasar population, and we provide further evidence to support this argument in the Supplementary Information and Extended Data Figs 9 and 10.

We compile a sample of 29 low-redshift AGNs with literature broad-line region size measurements from reverberation mapping[15], host stellar velocity dispersion (σ_*) measurements[23], and optical spectroscopy[24]. We use the well-established local $M_{\mathrm{BH}}-\sigma_*$ relation[25] to independently estimate black hole masses for the 29 AGNs. We supplement the 29 local AGNs with a sample of about 600 SDSS AGNs[26], where the host stellar velocity dispersion was estimated from the spectral decomposition of the SDSS spectrum into AGN and host galaxy components, and the broad-line region size R_{BLR} was estimated using the tight correlation between R_{BLR} and the AGN luminosity found in reverberation mapping studies[27]. We can then define a virial coefficient, $f \equiv GM_{\mathrm{BH}} \Big/ \left(R_{\mathrm{BLR}}\mathrm{FWHM}^2_{\mathrm{H}\beta}\right)$. At a given M_{BH}, f should not depend on $\mathrm{FWHM}_{\mathrm{H}\beta}$, if the latter is a faithful indicator of the broad-line region virial velocity. However, if $\mathrm{FWHM}_{\mathrm{H}\beta}$ is orientation-dependent, as suggested above, f will be anti-correlated with $\mathrm{FWHM}_{\mathrm{H}\beta}$.

Indeed, there is a strong dependence of f on $\mathrm{FWHM}_{\mathrm{H}\beta}$ at fixed M_{BH}, shown in Fig. 4, consistent with the orientation hypothesis. A direct consequence is that the standard virial black hole mass estimates using $\mathrm{FWHM}_{\mathrm{H}\beta}$ are subject to a significant uncertainty (about 0.4 dex), owing to this orientation dependence. To test this, we perform the same cross-correlation analysis as above, but for quasar subsamples divided by their virial black hole mass estimates based on $\mathrm{FWHM}_{\mathrm{H}\beta}$. The results are shown in Fig. 3b: there is no significant detection (1.64σ) in the clustering difference between the two quasar subsamples. This is in accordance

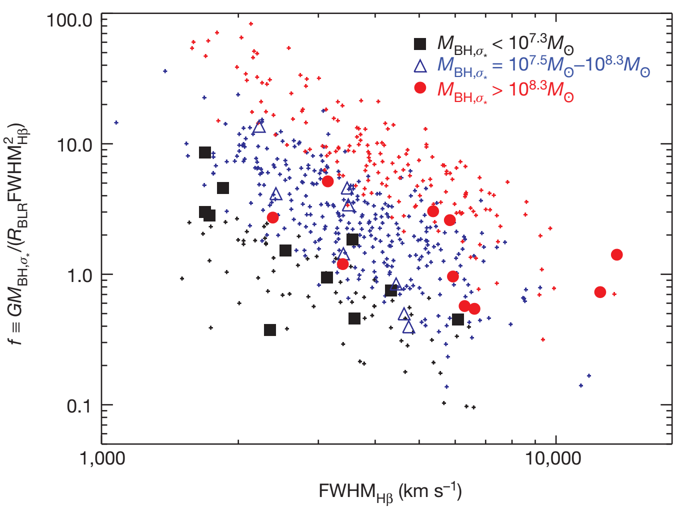


Figure 4 | The effect of orientation on $\mathrm{FWHM}_{\mathrm{H}\beta}$. The large symbols represent the 29 low-redshift AGNs that have both reverberation mapping data and host stellar velocity dispersion (σ_*) measurements. The small symbols represent a low-redshift SDSS AGN sample[26] with σ_* and AGN spectral measurements based on spectral decomposition. We use the stellar velocity dispersion measurements and the local relation between black hole mass and σ_* from inactive galaxies[25] to estimate the black hole mass (M_{BH,σ_*}) in these objects. We also estimate the average broad-line region size ($R_{\mathrm{BLR}} = c\tau$, where c is the speed of light, and τ is the measured reverberation mapping lag) in these objects, either from direct reverberation mapping measurements, or by using the tight correlation between the broad-line region size and AGN luminosity[27]. The ratio of M_{BH,σ_*} to $R_{\mathrm{BLR}}\mathrm{FWHM}^2_{\mathrm{H}\beta}/G$ (that is, the virial coefficient f) is plotted as a function of $\mathrm{FWHM}_{\mathrm{H}\beta}$, for different M_{BH,σ_*} values. The strong trends of f with $\mathrm{FWHM}_{\mathrm{H}\beta}$ at a given M_{BH,σ_*} suggest that the dispersion in $\mathrm{FWHM}_{\mathrm{H}\beta}$ does not reflect the underlying virial velocity of the broad-line region gas, and tends to bias the black hole mass estimates. This is in line with the fact that there is little vertical trend in the [O III] strength in the EV1 plane (Fig. 1).

with there being substantial overlap in the true black hole masses between the two subsamples, owing to the uncertainty in virial black hole mass estimates induced by using $\mathrm{FWHM}_{\mathrm{H}\beta}$. The division by $R_{\mathrm{Fe\,II}}$ provides a cleaner separation of high-mass black holes from low-mass black holes in our sample.

The collective evidence from this work leads to a simple interpretation of the observed main sequence of quasars (Fig. 1): the average Eddington ratio increases from left to right, and the dispersion in $\mathrm{FWHM}_{\mathrm{H}\beta}$ at fixed $R_{\mathrm{Fe\,II}}$ is largely an orientation effect. The many physical quasar properties correlated with EV1 are thus unified as being driven by changes in the average Eddington ratio of the black hole accretion. Although we do not discuss any physical model here, we suggest that the trends with the Eddington ratio are most probably caused by the systematic change in the shape of the accretion disk continuum and its interplay with the ambient emitting regions, which may in turn change the ionizing continuum (as seen by the emission-line regions) by modifying the structure of the accretion flow.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 March; accepted 23 July 2014.

1. Boroson, T. A. & Green, R. F. The emission-line properties of low-redshift quasi-stellar objects. *Astrophys. J.* **80** (suppl.), 109–135 (1992).
2. Sulentic, J. W., Zwitter, T., Marziani, P. & Dultzin-Hacyan, D. Eigenvector 1: an optimal correlation space for active galactic nuclei. *Astrophys. J.* **536,** L5–L9 (2000).
3. Wang, T., Brinkmann, W. & Bergeron, J. X-ray properties of active galactic nuclei with optical FeII emission. *Astron. Astrophys.* **309,** 81–96 (1996).
4. Boroson, T. A. Black hole mass and Eddington ratio as drivers for the observable properties of radio-loud and radio-quiet QSOs. *Astrophys. J.* **565,** 78–85 (2002).
5. Laor, A. The soft X-ray properties of a complete sample of optically selected quasars. II. Final results. *Astrophys. J.* **477,** 93–113 (1997).
6. Schneider, D. P. *et al.* The Sloan Digital Sky Survey quasar catalog. V. Seventh data release. *Astron. J.* **139,** 2360–2373 (2010).
7. Shen, Y. *et al.* A catalog of quasar properties from Sloan Digital Sky Survey data release 7. *Astrophys. J.* **194** (suppl.), 45 (2011).
8. Stern, J. & Laor, A. Type 1 AGN at low z—III. The optical narrow-line ratios. *Mon. Not. R. Astron. Soc.* **431,** 836–857 (2013).
9. Zhang, K., Wang, T.-G., Gaskell, C. M. & Dong, X.-B. The Baldwin effect in the narrow emission lines of active galactic nuclei. *Astrophys. J.* **762,** 51 (2013).
10. Baldwin, J. A. Luminosity indicators in the spectra of quasi-stellar objects. *Astrophys. J.* **214,** 679–684 (1977).
11. Dopita, M. A. & Sutherland, R. S. Spectral signatures of fast shocks. II. Optical diagnostic diagrams. *Astrophys. J.* **455,** 468–479 (1995).
12. Wright, E. L. *et al.* The Wide-field Infrared Survey Explorer (WISE): mission description and initial on-orbit performance. *Astron. J.* **140,** 1868–1881 (2010).
13. Antonucci, R. Unified models for active galactic nuclei and quasars. *Annu. Rev. Astron. Astrophys.* **31,** 473–521 (1993).
14. Urry, C. M. & Padovani, P. Unified schemes for radio-loud active galactic nuclei. *Publ. Astron. Soc. Pacif.* **107,** 803–845 (1995).
15. Peterson, B. M. *et al.* Central masses and broad-line region sizes of active galactic nuclei. II. A homogeneous analysis of a large reverberation-mapping database. *Astrophys. J.* **613,** 682–699 (2004).
16. Laor, A. On black hole masses and radio loudness in active galactic nuclei. *Astrophys. J.* **543,** L111–L114 (2000).
17. Vestergaard, M. & Peterson, B. M. Determining central black hole masses in distant active galaxies and quasars. II. Improved optical and UV scaling relationships. *Astrophys. J.* **641,** 689–709 (2006).
18. Bardeen, J. M., Bond, J. R., Kaiser, N. & Szalay, A. S. The statistics of peaks of Gaussian random fields. *Astrophys. J.* **304,** 15–61 (1986).
19. Ahn, C. *et al.* The tenth data release of the Sloan Digital Sky Survey: first spectroscopic data from the SDSS-III Apache Point Observatory Galactic Evolution Experiment. *Astrophys. J.* **211** (Supp.), 17 (2014).
20. Shen, Y. *et al.* Cross-correlation of SDSS DR7 quasars and DR10 BOSS galaxies: the weak luminosity dependence of quasar clustering at $z \sim 0.5$. *Astrophys. J.* **778,** 98 (2013).
21. Wills, B. J. & Browne, I. W. A. Relativistic beaming and quasar emission lines. *Astrophys. J.* **302,** 56–63 (1986).
22. Runnoe, J. C., Brotherton, M. S., Shang, Z., Wills, B. J. & DiPompeo, M. A. The orientation dependence of quasar single-epoch black hole mass scaling relationships. *Mon. Not. R. Astron. Soc.* **429,** 135–149 (2013).
23. Park, D., Kelly, B. C., Woo, J.-H. & Treu, T. Recalibration of the virial factor and M_{BH}-σ_* relation for local active galaxies. *Astrophys. J.* **203** (suppl.), 6 (2012).
24. Marziani, P. *et al.* An optical spectroscopic atlas of low-redshift active galactic nuclei. *Astrophys. J.* **145** (suppl.), 199–211 (2003).
25. Kormendy, J. & Ho, L. C. Coevolution (or not) of supermassive black holes and host galaxies. *Annu. Rev. Astron. Astrophys.* **51,** 511–653 (2013).

26. Shen, J., Vanden Berk, D. E., Schneider, D. P. & Hall, P. B. The black hole-bulge relationship in luminous broad-line active galactic nuclei and host galaxies. *Astron. J.* **135,** 928–946 (2008).
27. Bentz, M. C. *et al.* The radius-luminosity relationship for active galactic nuclei: the effect of host-galaxy starlight on luminosity measurements. II. The full sample of reverberation-mapped AGNs. *Astrophys. J.* **697,** 160–181 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements Support for the work of Y.S. was provided by NASA through Hubble Fellowship grant number HST-HF-51314.01, awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy for NASA, under contract number NAS 5-26555. L.C.H. acknowledges support from the Kavli Foundation, Peking University, and the Chinese Academy of Science through grant number XDB09030102 (Emergence of Cosmological Structures) from the Strategic Priority Research Program. This work makes extensive use of SDSS-I/II and SDSS-III data (http://www.sdss.org/ and http://www.sdss3.org/).

Author Contributions Y.S. and L.C.H. co-developed the idea. Y.S. performed the measurements and analysis. Both authors contributed to the interpretation and manuscript writing.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.S. (yshen@obs.carnegiescience.edu).

doi:10.1038/nature13734

Probing excitonic dark states in single-layer tungsten disulphide

Ziliang Ye[1]*, Ting Cao[2,3]*, Kevin O'Brien[1], Hanyu Zhu[1], Xiaobo Yin[1], Yuan Wang[1], Steven G. Louie[2,3] & Xiang Zhang[1,3,4,5]

Transition metal dichalcogenide (TMDC) monolayers have recently emerged as an important class of two-dimensional semiconductors with potential for electronic and optoelectronic devices[1,2]. Unlike semimetallic graphene, layered TMDCs have a sizeable bandgap[3]. More interestingly, when thinned down to a monolayer, TMDCs transform from indirect-bandgap to direct-bandgap semiconductors[4,5], exhibiting a number of intriguing optical phenomena such as valley-selective circular dichroism[6–8], doping-dependent charged excitons[9,10] and strong photocurrent responses[11]. However, the fundamental mechanism underlying such a strong light–matter interaction is still under intensive investigation. First-principles calculations have predicted a quasiparticle bandgap much larger than the measured optical gap, and an optical response dominated by excitonic effects[12–14]. In particular, a recent study based on a GW plus Bethe–Salpeter equation (GW-BSE) approach, which employed many-body Green's-function methodology to address electron–electron and electron–hole interactions, theoretically predicted a diversity of strongly bound excitons[14]. Here we report experimental evidence of a series of excitonic dark states in single-layer WS_2 using two-photon excitation spectroscopy. In combination with GW-BSE theory, we prove that the excitons are of Wannier type, meaning that each exciton wavefunction extends over multiple unit cells, but with extraordinarily large binding energy (~0.7 electronvolts), leading to a quasiparticle bandgap of 2.7 electronvolts. These strongly bound exciton states are observed to be stable even at room temperature. We reveal an exciton series that deviates substantially from hydrogen models, with a novel energy dependence on the orbital angular momentum. These excitonic energy levels are experimentally found to be robust against environmental perturbations. The discovery of excitonic dark states and exceptionally large binding energy not only sheds light on the importance of many-electron effects in this two-dimensional gapped system, but also holds potential for the device application of TMDC monolayers and their heterostructures[15] in computing, communication and bio-sensing.

An exciton is a bound state formed by an excited electron and hole owing to the Coulomb attraction between these two quasiparticles[16]. Such bound states often play an important role in the optical properties of low-dimensional materials[17], owing to their strong spatial confinement and reduced screening effect compared to bulk solids. In a two-dimensional (2D) gapped system with dipole-allowed interband transitions, the optical absorption spectrum in the non-interacting limit exhibits a step function. Strong electron–hole interaction redshifts a large amount of the spectral weight, resulting in a qualitatively different spectrum with a series of new excitonic levels below the quasiparticle bandgap. In quasi-2D quantum wells, the electron–hole interaction is weak. Therefore, by measuring the energy difference between the first excitonic peak and band-edge absorption step, the exciton binding energy can be unambiguously determined; it usually has an energy of tens of meV and is vulnerable to environment screening and temperature broadening. However, recent experiments on a single-layer TMDC like MoS_2 found no absorption step[4,5]. Instead, two absorption peaks from spin–orbit splitting were detected[4,5] around the Kohn–Sham bandgap energy, as given by density functional theory (DFT) within the local density approximation. The peaks were initially interpreted as direct band edge transitions. In sharp contrast, more accurate first-principles calculations on MoS_2 monolayer using the GW method[18] predicted a quasiparticle bandgap that was larger than the initial experimental reported value by nearly one electronvolt[12–14]. Relevant calculations based on first-principles GW-BSE theory[19] showed this energy gap discrepancy to originate in strong excitonic effects. It is therefore critical to uncover the underlying physics of the strong light–matter interaction in such a 2D system.

We probed the excitonic effects in monolayer WS_2, also an important TMDC material, using two-photon excitation spectroscopy[20]. At the simplest level, if an electron–hole pair interacts through a Coulomb attractive central potential, it will form a series of excitonic Rydberg-like states with definite parity, similar to the hydrogen model. For WS_2, the breaking of rotational and inversion symmetry owing to the crystal structure and the spatial dependence of screening will modify the energy and symmetry of the states from those of the 2D Rydberg series. However, for exciton states with an electron–hole wavefunction that is large compared to the unit cell size (as shown below for WS_2), specific parity may still be assigned to each excitonic state. Incident photons can excite the electronic system from the ground state to one of these excitonic states (Fig. 1a). In addition to energy conservation, the selection rule of such a transition depends on the symmetry of the final state: for systems with dipole-allowed interband transitions (which is the case for WS_2), one-photon transitions can only reach excitonic states with even parity, while two-photon transitions reach states with odd parity. The two-photon resonances are also known as excitonic dark states as they do not appear in the linear optical spectrum. These dark states are good gauges for excitonic effects, since there is little impurity and bandgap absorption background in the two-photon spectrum. Owing to the direct bandgap in this WS_2 monolayer, we monitor the two-photon absorption induced luminescence (which we abbreviate to two-photon luminescence, TPL) with a high signal-to-noise ratio. The luminescence results from the radiative recombination of the excitonic ground state, following the rapid non-radiative relaxation from the two-photon excited excitonic dark states to the exciton ground state (Fig. 1a). By scanning the excitation laser energy, we obtain a complete two-photon spectrum, assuming the relaxation and emission efficiency are independent of the excitation energy.

Our samples are WS_2 monolayers directly exfoliated on fused quartz substrates. A typical light emission spectrum is shown in Fig. 1b, excited by an ultrafast laser (pulses of 190 fs duration) at a wavelength of 990 nm (1.25 eV) at a sample temperature of 10 K. The two peaks observed at 2.0 eV and 2.04 eV correspond to the exciton and negatively charged trion emissions from the direct bandgap at K and K′ valleys in the Brillouin zone, consistent with the absorption peaks in the reflectance spectrum (Supplementary Information section S1) The emitted photon energies of both peaks are much higher than those of the excitation photon, and

[1]NSF Nano-scale Science and Engineering Center (NSEC), 3112 Etcheverry Hall, University of California, Berkeley, California 94720, USA. [2]Department of Physics, University of California, Berkeley, California 94720, USA. [3]Material Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA. [4]Department of Physics, King Abdulaziz University, Jeddah 21589, Saudi Arabia. [5]Kavli Energy NanoSciences Institute at the University of California, Berkeley, and Lawrence Berkeley National Laboratory, Berkeley, California 94704, USA.
*These authors contributed equally to this work.

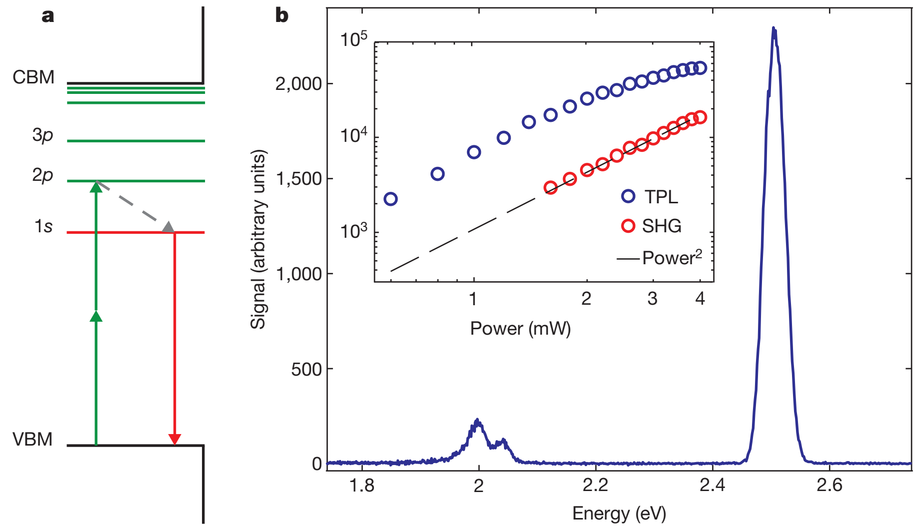


Figure 1 | Probing the dark exciton states in single-layer WS_2 by two-photon luminescence. **a**, Schematic of the two-photon luminescence (TPL) process in single-layer WS_2. Under two-photon excitation, electrons transition to one of the excitonic dark states with odd parity (double green arrow). Following the excitation, the exciton experiences a fast relaxation to the excitonic ground state (grey arrow) and emits a photon (red arrow). The two-photon selection rule exclusively eliminates the one-photon transition background and reveals the excitonic excited states. States are labelled *s* (red) or *p* (green) according to the excitonic envelope wavefunction character. CBM and VBM represent respectively the conduction band minimum and the valence band maximum. **b**, Main panel: measured WS_2 emission spectrum excited by an ultrafast pulsed laser at 10 K. The peaks at 2.04 eV and 2 eV are the A exciton (1*s* state) and its trion peak, respectively. The lower-energy peak is stronger than the higher-energy one due to the exciton–trion equilibrium reached during the emission stage at low temperature. The excitation pulse is at 1.25 eV with a pulse width of about 190 ± 20 fs, which results in the 2.5 eV peak as the SHG signal. Inset, the power dependence of the SHG and TPL signals. At a low excitation level, both of them exhibit quadratic power dependence, confirming the two-photon absorption nature of the luminescence, until the TPL signal saturates at a high excitation level. The TPL signal represents the amplitude of the trion peak.

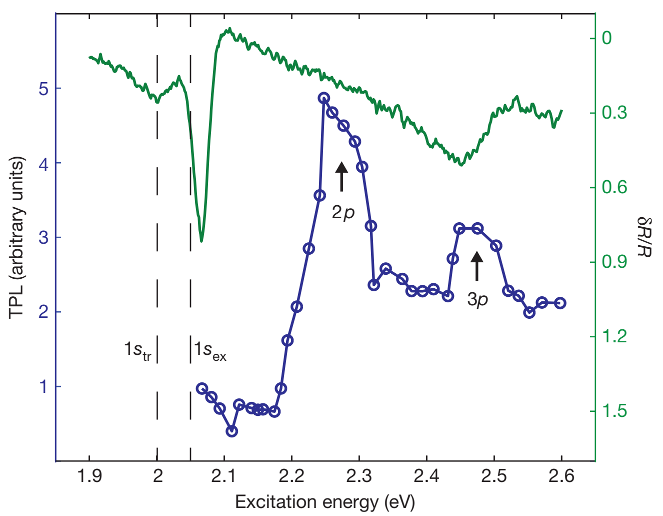


Figure 2 | Extraordinarily strong excitonic effect in monolayer WS_2. Two-photon absorption (blue) and one-photon absorption (green) spectra are measured in single-layer WS_2 at 10 K. In the two-photon absorption spectrum, 2*p* and 3*p* resonances are observed at 2.28 eV and 2.48 eV, respectively, on top of a plateau background. For comparison, the one-photon absorption spectrum, measured as the relative reflectance signal ($\delta R/R$), exhibits no corresponding features except a B exciton (1*s*) related absorption resonance at 2.45 eV. Additionally, the A exciton ($1s_{ex}$) and trion ($1s_{tr}$) absorption peaks are detected consistently with the TPL emission peaks (Fig. 1b), with a 20 meV Stoke shift, and are marked at 2.04 and 2 eV, respectively, by black dashed lines. The energy difference between the A exciton 1*s* state emission peak and the 3*p* state absorption peak is 0.44 eV, which yields the lower bound for the exciton binding energy in monolayer WS_2. This binding energy is extraordinarily large for a Wannier exciton, and implies a dominating excitonic mechanism for the intense light–matter interaction in 2D TMDCs. The total excitation scan is achieved by tuning an output beam of an optical parametric oscillator over a 600 meV span, with a scanning resolution of about 15 meV (Supplementary Information section S3). Similar results are repeated in more than 5 flakes.

therefore they can only originate from TPL. The peak at 2.5 eV is the second harmonic generation (SHG) emission. The two-photon origin of these emissions is further confirmed in Fig. 1b inset. Both the TPL and SHG signals show a quadratic power dependence, suggesting that the emission is indeed induced by two-photon absorption. The TPL saturates at higher power as a consequence of heating or exciton–exciton annihilation effects[21,22]. The trion peak amplitude is selected as our TPL signal.

We collect the TPL signal, while scanning the excitation laser energy from 2.05 to 2.6 eV, to acquire the full two-photon spectrum. We observed two important resonances of similar linewidths in the two-photon spectrum, occurring at 2.28 and 2.48 eV, corresponding to two excitonic dark excited states (Fig. 2). The absorption spectrum of a WS_2 monolayer is plotted for comparison, where the A exciton (the 1*s* state) and its trion result in two absorption peaks at 2.04 eV and 2 eV, respectively. Near these one-photon resonances, TPL is negligible, consistent with the 1*s* nature of these states. On the other hand, no significant one-photon absorption is observed near the excitonic dark states, except for the B exciton (the other 1*s* state) at 2.45 eV which results from the spin–orbit splitting in the valence band. Such a complementary feature reflects the symmetry of the observed excitonic states. Hence, we label the TPL peaks as the 2*p* and 3*p* state of the A exciton series. Accordingly, the 1*s*–2*p* and 1*s*–3*p* separations are 0.24 eV and 0.44 eV, respectively. The extraordinarily large 1*s*–*np* ($n =$ 2, 3) separations suggest that the exciton binding energy, defined as the separation between the 1*s* exciton ground state and the conduction band edge, is larger than 0.44 eV, which also indicates a significant self-energy contribution to the quasi-particle bandgap. Our discovery demonstrates that the previously claimed band-to-band transition mechanism in the optical response of monolayer WS_2 is inaccurate, as we show here that the optical response is dominated by excitonic states within the bandgap, in agreement with the GW-BSE calculation of MoS_2 (ref. 14). The real quasiparticle bandgap is much larger than previously reported. This finding is expected be general for other TMDC monolayers of similar structure.

We used the *ab initio* GW method[18] to calculate the quasiparticle band structure and the *ab initio* GW-BSE approach[19] to calculate the excitonic states and optical spectrum of a WS_2 monolayer (Fig. 3a), employing the BerkeleyGW package[23]. The principal and orbital quantum numbers of

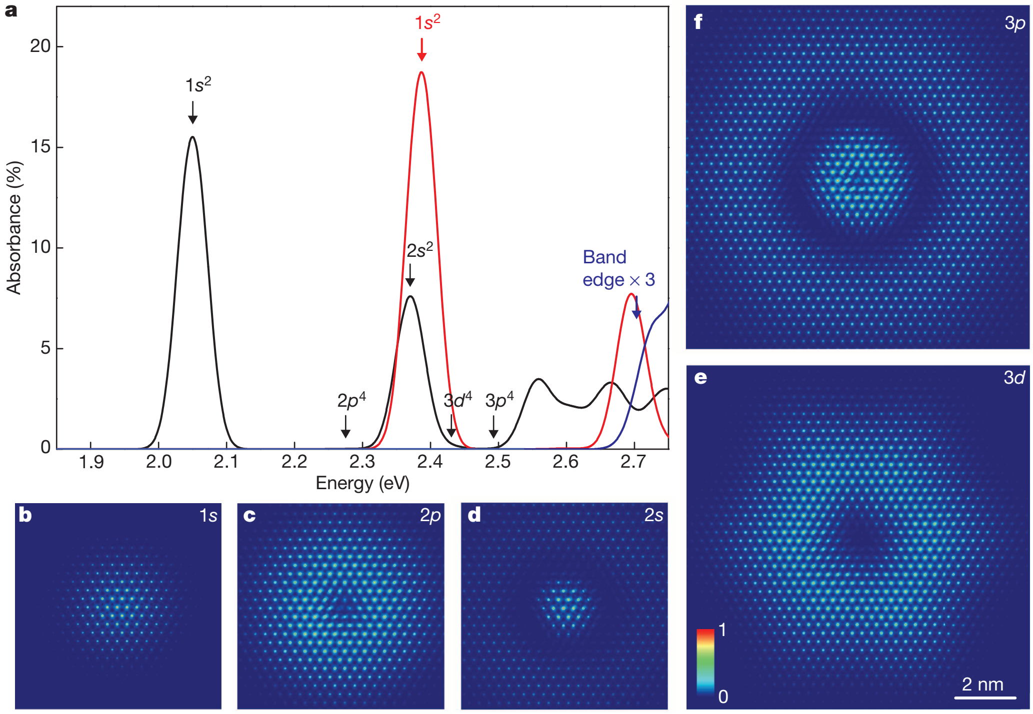


Figure 3 | One-photon absorption spectra and real-space exciton wavefunctions in monolayer WS_2 from *ab initio* GW-BSE calculations. **a**, The optical absorption of the A (black) and B (red) exciton series considering electron–hole interaction. The blue curve is the optical absorption spectrum, obtained without considering electron–hole interaction, where the quasiparticle bandgap is about 2.7 eV (blue arrow). The excitonic states of A and B exciton series, with electron–hole interaction included, are calculated (shown in **b**–**f**, see below) and labelled (in **a**) by black and red arrows, respectively, up to 2.5 eV. The computed 1*s*, 2*p* and 3*p* states of the A exciton are at 2.05 eV, 2.28 eV and 2.49 eV, respectively, and are in excellent agreement with the experimental measurements. Although the orbital notation of a 2D hydrogen atom is adopted to label the exciton states, the excitonic series significantly deviates from a hydrogenic series, as discussed in the main text. The degeneracy labels in the superscript include both the degeneracy of valleys and orbital angular momentum. **b**–**f**, The plots are modulus squared of the real-space exciton wavefunction projected onto the WS_2 plane, with the hole position fixed near a W atom at the centre of the plot. These wavefunctions share similar in-plane nodal structures with the excited states in a hydrogen atom, and therefore enable the eigenstates to be labelled with a principal and an orbital quantum number. The Wannier nature of the excitons is clear, with the radii much larger than the unit cell. The colour scale is the normalized wavefunction probability and applies to panels **b**–**f**.

each exciton state are identified by analysing the character of the exciton's real-space wavefunction (Fig. 3b–f). Specifically, the nodal characters along the radial direction are unique for each exciton state and have a one-to-one correspondence with those of the 2D Rydberg series. Consistent with the selection rule of one-photon absorption for dipole-allowed materials, we find that the '*s*' state is one-photon active or bright, while the other ('*p*' and '*d*') excitons are one-photon inactive or dark (see detailed analysis in Supplementary Information section S2). Clearly, the calculated 2*p* and 3*p* states, marked at 2.28 and 2.49 eV in Fig. 3a, agree well with the experimental results, which confirms our observation of dark excitonic states in WS_2 monolayer. The calculated positions of the 1*s* state of the A exciton series (2.04 eV) and B exciton series (2.4 eV) also agree well with the experimental spectrum. As is evident from the real-space wavefunctions in Fig. 3b–f, the excitons in monolayer WS_2 have a Wannier nature, with their in-plane radii much larger than the unit cell dimension. As mentioned above, owing to the broken inversion symmetry of the TMDC monolayer, the linear absorption selection rule is not exact. The exciton *p* states acquire a small but finite oscillator strength in our calculation, with the oscillator strength two orders of magnitude smaller than that of the *s* state in the same shell.

In spite of its Wannier character, we found that the exciton series in monolayer WS_2 deviates significantly from a 2D hydrogen model. Much smaller splitting between 1*s* and other excited states is observed, in accordance with recent GW-BSE calculations[14] (see detailed comparisons in Supplementary Information section S4). In addition, in a hydrogen atom, orbitals with the same principal quantum number are degenerate. However, for the WS_2 excitons, our calculations show that states in the same shell but of higher orbital angular momentums are at lower energy levels, that is, $E_{3d} < E_{3p} < E_{3s}$. These two exotic energy-level behaviours are caused by a strong spatial-dependent dielectric screening: in an atomically thin semiconductor, the screening effect at more than a certain distance is weaker when the separation between the electron and hole is bigger, which is known as the anti-screening effect in 1D carbon nanotubes[24] and as the dielectric confinement effect in 2D quantum wells[25]. Since the wavefunction of excitonic states with higher principal or higher orbital quantum number features a larger nodal structure near the hole (that is, a larger average electron–hole separation), weaker screening at larger separation leads to enhanced Coulomb attraction in the excited states and therefore a lowering of their excitation energies as compared with those of the hydrogen model[24]. Also, because of the degeneracy of the K and K′ valleys in the TMDC system, each *s* level has two degenerate states, while each *p* and *d* level has four degenerate states if perfect rotational symmetry is assumed. All of these features are expected to be quite general for 2D TMDC excitons.

The GW quasiparticle bandgap is calculated to be ~2.7 eV, indicated by the blue arrow in Fig. 3. Comparing this with the 1*s* exciton energy found in either our experiments or our GW-BSE calculations, we obtain an exciton binding energy of ~0.7 eV. Such an exceptionally large binding energy is more than ten times that found for the excitons in bulk WS_2 (ref. 3) and other traditional bulk semiconductors such as Si and GaAs (ref. 16), and comparable to those found for excitons in carbon nanotubes[20,26]. The large binding energy results from the combined effects of reduced dimensionality, relatively large effective masses and weak dielectric screening, which renders the excitons observable even at room temperature. Similar effects were also found in carbon nanotubes and inorganic-organic hybrid perovskites[20,27].

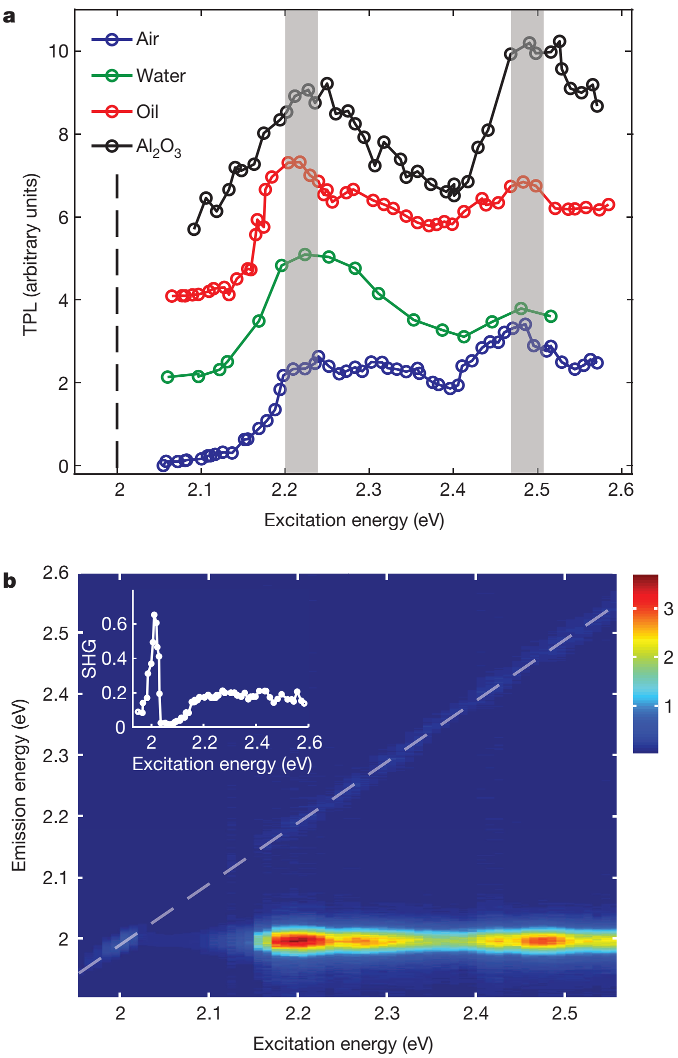


Figure 4 | Excitonic energy levels are robust to changes in the dielectric environment and to temperature changes. a, Room-temperature two-photon spectra of single-layer WS_2 with different top capping layers that tune the dielectric environment immediately adjacent to the atomic layer. The curves respectively represent the uncapped ($\varepsilon_{ave} = 1.625$, where ε_{ave} is the average dielectric constant between capping layers and the substrate), water capped ($\varepsilon_{ave} = 1.97$), immersion-oil capped ($\varepsilon_{ave} = 2.25$) and Al_2O_3 capped ($\varepsilon_{ave} = 2.57$) samples, and each curve is adjusted to a similar vertical scale and shifted for better visualization. The emission peak is at 2 eV, marked by the vertical black dashed line. Evidently, the 2*p* and 3*p* peak positions remain roughly unchanged within experimental error, marked by the grey bands at 2.22 ± 0.02 eV and 2.49 ± 0.02 eV, respectively. Therefore, the 1*s*–*np* ($n = 2, 3$) separation is approximately the same as the low-temperature uncapped result (Fig. 2), suggesting that the excitation energy of the low-energy exciton levels are relatively insensitive to dielectric environmental and temperature perturbations, as discussed in the main text. **b**, Main panel: measured emission spectra at different excitation energies of an immersion-oil capped WS_2 monolayer at room temperature. The horizontal line signal is the TPL emission, with two hotspots along the line corresponding to the 2*p* and 3*p* two-photon absorption peaks. Colour scale represents the normalized emission intensity. The SHG signal due to the broken inversion symmetry in the monolayer is observed (along the dashed line as an eye guide). At the intersection between the SHG and TPL line, the SHG signal experiences an excitonic enhancement from the A exciton 1*s* state (inset).

The excitonic ground state and low-energy excited states with large binding energy are robust to environmental perturbations owing to the opposite effects of the dielectric screening on the exciton binding energy and the quasiparticle self-energy[20,28]. We demonstrate this by measuring two-photon spectra of monolayer WS_2 with different dielectric capping layers, including water, immersion oil and aluminium oxide; the average dielectric constants of these capping layers at optical frequency range from 1.7 to 2.5. In all capped samples, we observed the 2*p* and 3*p* resonances even at room temperature (Fig. 4a). We find no significant shift in the excitation energy of either the *s* or the *p* states with different capping layers, except for an overall temperature-related redshift (0.04 eV) and linewidth broadening compared with measurements at 10 K (Fig. 2). The 1*s*–2*p* and 1*s*–3*p* energy differences remain roughly unchanged, ~0.2 and 0.5 eV, respectively. This robustness indicates that the measured excitation energies for the 2*p* and 3*p* states are intrinsic to the monolayer, thus agreeing well with those from an *ab initio* GW-BSE calculation for the vacuum condition. Together with the TPL signal, SHG is also observed as a slanted straight line in the excitation-emission spectra (Fig. 4b). At room temperature, the exciton–trion separation is no longer distinguishable, but the 2*p* and 3*p* absorption peaks remain prominent. An SHG resonance occurs as the TPL and SHG lines cross each other, and this resonance is known as the exciton enhanced SHG effect[29].

We have experimentally revealed 2D excitonic dark states in a WS_2 monolayer. These observations unveil an intense many-electron effect in this class of 2D gapped systems. The determined bandgap size would allow us to accurately design heterostructures consisting of a TMDC monolayer and other materials. Our discovery of extraordinarily strong excitons in a TMDC provides a basis for exploiting the unusual light–matter interactions resulting from strong many-electron effects, and should also help the development of emerging 2D electronic and optoelectronic applications.

Received 29 January; accepted 31 July 2014.
Published online 27 August 2014.

1. Wang, Q. H., Kalantar-Zadeh, K., Kis, A. & Coleman, J. N. Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. *Nature Nanotechnol.* **7,** 699–712 (2012).
2. Radisavljevic, B., Radenovic, A., Brivio, J., Giacometti, V. & Kis, A. Single-layer MoS_2 transistors. *Nature Nanotechnol.* **6,** 147–150 (2011).
3. Beal, A. R., Knights, J. C. & Liang, W. Y. Transmission spectra of some transition metal dichalcogenides. II. Group VIA: trigonal prismatic coordination. *J. Phys. C* **5,** 3540–3551 (1972).
4. Mak, K. F., Lee, C., Hone, J., Shan, J. & Heinz, T. F. Atomically thin MoS_2: a new direct-gap semiconductor. *Phys. Rev. Lett.* **105,** 136805 (2010).
5. Splendiani, A. *et al.* Emerging photoluminescence in monolayer MoS_2. *Nano Lett.* **10,** 1271–1275 (2010).
6. Zeng, H., Dai, J., Yao, W., Xiao, D. & Cui, X. Valley polarization in MoS_2 monolayers by optical pumping. *Nature Nanotechnol.* **7,** 490–493 (2012).
7. Mak, K. F., He, K., Shan, J. & Heinz, T. F. Control of valley polarization in monolayer MoS_2 by optical helicity. *Nature Nanotechnol.* **7,** 494–498 (2012).
8. Cao, T. *et al.* Valley-selective circular dichroism of monolayer molybdenum disulphide. *Nature Commun.* **3,** 887 (2012).
9. Mak, K. F. *et al.* Tightly bound trions in monolayer MoS_2. *Nature Mater.* **12,** 207–211 (2012).
10. Ross, J. S. *et al.* Electrical control of neutral and charged excitons in a monolayer semiconductor. *Nature Commun.* **4,** 1474 (2013).
11. Britnell, L. *et al.* Strong light-matter interactions in heterostructures of atomically thin films. *Science* **340,** 1311–1314 (2013).
12. Ramasubramaniam, A. Large excitonic effects in monolayers of molybdenum and tungsten dichalcogenides. *Phys. Rev. B* **86,** 115409 (2012).
13. Cheiwchanchamnangij, T. & Lambrecht, W. R. L. Quasiparticle band structure calculation of monolayer, bilayer, and bulk MoS_2. *Phys. Rev. B* **85,** 205302 (2012).
14. Qiu, D. Y., Felipe, H. & Louie, S. G. Optical spectrum of MoS_2: many-body effects and diversity of exciton states. *Phys. Rev. Lett.* **111,** 216805 (2013).
15. Geim, A. K. & Grigorieva, I. V. Van der Waals heterostructures. *Nature* **499,** 419–425 (2013).
16. Knox, R. S. *Theory of Excitons* (Academic, 1963).
17. Scholes, G. D. & Rumbles, G. Excitons in nanoscale systems. *Nature Mater.* **5,** 683–696 (2006).
18. Hybertsen, M. S. & Louie, S. G. Electron correlation in semiconductors and insulators: band gaps and quasiparticle energies. *Phys. Rev. B* **34,** 5390–5413 (1986).
19. Rohlfing, M. & Louie, S. G. Electron-hole excitations and optical spectra from first principles. *Phys. Rev. B* **62,** 4927–4944 (2000).
20. Wang, F., Dukovic, G., Brus, L. E. & Heinz, T. F. The optical resonances in carbon nanotubes arise from excitons. *Science* **308,** 838–841 (2005).
21. Ye, Y. *et al.* Exciton-dominant electroluminescence from a diode of monolayer MoS_2. *Appl. Phys. Lett.* **104,** 193508 (2014).
22. Kumar, N. *et al.* Exciton-exciton annihilation in $MoSe_2$ monolayers. *Phys. Rev. B* **89,** 125427 (2014).
23. Deslippe, J. *et al.* BerkeleyGW: A massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures. *Comput. Phys. Commun.* **183,** 1269–1289 (2012).

24. Deslippe, J. *et al.* Electron-hole interaction in carbon nanotubes: novel screening and exciton excitation spectra. *Nano Lett.* **9,** 1330–1334 (2009).
25. Keldysh, L. V. Coulomb interaction in thin semiconductor and semimetal films. *J. Exp. Theor. Phys. Lett.* **29,** 658–660 (1979).
26. Spataru, C. D., Ismail-Beigi, S., Benedict, L. X. & Louie, S. G. Excitonic effects and optical spectra of single-walled carbon nanotubes. *Phys. Rev. Lett.* **92,** 077402 (2004).
27. Ishihara, T., Takahashi, J. & Goto, T. Optical properties due to electronic transitions in two-dimensional semiconductors $(C_nH_{2n+1}NH_3)_2PbI_4$. *Phys. Rev. B* **42,** 11099–11107 (1990).
28. Maultzsch, J. *et al.* Exciton binding energies in carbon nanotubes from two-photon photoluminescence. *Phys. Rev. B* **72,** 241402 (2005).
29. Malard, L. M., Alencar, T. V., Barboza, A. P. M., Mak, K. F. & de Paula, A. M. Observation of intense second harmonic generation from MoS_2 atomic crystals. *Phys. Rev. B* **87,** 201401 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements This research was supported by the Lawrence Berkeley National Laboratory through the Office of Basic Energy Sciences, US Department of Energy under contract no. DE-AC02-05CH11231: the experimental work was supported by Sub-wavelength Metamaterial Design, Physics and Applications Program, and the theory part was supported by the Theory Program (GW-BSE calculations and simulations) and by the SciDAC Program on Excited State Phenomena (computer codes and algorithm developments), with computer time provided by the DOE National Energy Research Scientific Computing Center (NERSC). Z.Y. acknowledges discussions with T. Ishihara and F. Wang.

Author Contributions Z.Y. and X.Z. initiated this research topic; Z.Y., K.O., X.Y. and Y.W. conducted the optical experiments; Z.Y. and H.Z. prepared samples; T.C. and S.G.L. performed the first-principles calculations; Z.Y., T.C., S.G.L. and X.Z. analysed the results and prepared the manuscript. All authors contributed to discussions and manuscript revision.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to X.Z. (xzhang@me.berkeley.edu) and S.G.L. (sglouie@berkeley.edu).

doi:10.1038/nature13721

Observational evidence for interhemispheric hydroxyl-radical parity

P. K. Patra[1,2], M. C. Krol[3], S. A. Montzka[4], T. Arnold[5], E. L. Atlas[6], B. R. Lintner[7], B. B. Stephens[8], B. Xiang[9], J. W. Elkins[4], P. J. Fraser[10], A. Ghosh[1,11], E. J. Hintsa[4,12], D. F. Hurst[4,12], K. Ishijima[1], P. B. Krummel[10], B. R. Miller[4,12], K. Miyazaki[1], F. L. Moore[4,12], J. Mühle[5], S. O'Doherty[13], R. G. Prinn[14], L. P. Steele[10], M. Takigawa[1], H. J. Wang[15], R. F. Weiss[5], S. C. Wofsy[9] & D. Young[13]

The hydroxyl radical (OH) is a key oxidant involved in the removal of air pollutants and greenhouse gases from the atmosphere[1–3]. The ratio of Northern Hemispheric to Southern Hemispheric (NH/SH) OH concentration is important for our understanding of emission estimates of atmospheric species such as nitrogen oxides and methane[4–6]. It remains poorly constrained, however, with a range of estimates from 0.85 to 1.4 (refs 4, 7–10). Here we determine the NH/SH ratio of OH with the help of methyl chloroform data (a proxy for OH concentrations) and an atmospheric transport model that accurately describes interhemispheric transport and modelled emissions. We find that for the years 2004–2011 the model predicts an annual mean NH–SH gradient of methyl chloroform that is a tight linear function of the modelled NH/SH ratio in annual mean OH. We estimate a NH/SH OH ratio of 0.97 ± 0.12 during this time period by optimizing global total emissions and mean OH abundance to fit methyl chloroform data from two surface-measurement networks and aircraft campaigns[11–13]. Our findings suggest that top-down emission estimates of reactive species such as nitrogen oxides in key emitting countries in the NH that are based on a NH/SH OH ratio larger than 1 may be overestimated.

As the primary atmospheric oxidant, the OH radical has a key role in the removal or production of major air pollutants, greenhouse gases and many ozone-depleting substances[1–3]. A better understanding of the NH/SH OH ratio will lead to significantly improved source estimates of reactive species and to an improved prediction of chemistry–climate interactions from changing human activities. Because of its very short lifetime (~1 s), OH manifests high spatiotemporal variability. Moreover, because OH concentrations are typically very low (~10^6 molecules cm^{-3}), *in situ* measurements are challenging, and large differences between observations have prevented a direct validation of model-simulated OH distributions or an evaluation of uncertainties in the chemical mechanisms responsible for OH recycling under different environmental conditions[14–17]. Rather, indirect estimates of the total abundance and interannual variations of global OH have been made with methyl chloroform (CH_3CCl_3) or ^{14}CO measurements and simulations by chemistry-transport models (CTMs)[8–10,12,18–20].

Constraints on the meridional OH gradient are needed for estimating hemispheric source and sink magnitudes of gases and aerosols that are produced or destroyed through reactions with OH (see Methods). For example, large increases in NO_x ($NO + NO_2$) emissions from China compared with emission inventories are estimated by using a data assimilation system[5] and the CHASER (Chemical AGCM for Studies of Atmospheric Environment and Radiative forcing[21]) CTM. An overestimate of NH OH could account for such discrepancies, because it affects the hemispheric budgets of NO_x and other reactive species such as carbon monoxide (CO) and methane (CH_4). CHASER-simulated OH is ~26% higher in the NH than in the SH, and this model predicts a smaller than observed NH–SH CH_4 gradient[6]. In the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), the multi-model average and 1σ spread of annual mean NH/SH OH ratios was 1.28 ± 0.10 (range 1.13–1.42). The simulated NH/SH OH ratio being greater than 1 is related primarily to OH production from modelled ozone, which is biased high in the NH and low in the SH, compared with observations[4]. Other evidence exists for significantly lower NH/SH OH ratios. On the basis of ^{14}CO and CH_3CCl_3 observations, the NH/SH ratio of OH is suggested to be significantly lower than 1 (refs 7, 8). A NH/SH OH ratio of 0.98 using models has been shown[9], but it was concluded that accurate estimation of the ratio would require more accurate CH_3CCl_3 emissions (also in ref. 8) and models that accurately describe interhemispheric exchange.

We constrain the NH/SH ratio of OH by using an atmospheric general circulation model (AGCM)-based CTM, the Japan Agency for Marine-Earth Sciences and Technology (JAMSTEC) Atmospheric Chemistry Transport Model (ACTM)[6,22]. CH_3CCl_3 simulations are performed with two spatially distinct tropospheric OH distributions, namely, ACTM_0.99 and ACTM_1.26, which have annual mean NH/SH OH ratios of 0.99 and 1.26, respectively[21,23] (Extended Data Fig. 1), with the NH and SH separated at the geographical Equator. The NH/SH OH ratios are 0.87 for ACTM_0.99 and 1.18 for ACTM_1.26, when the two hemispheres are divided in accordance with the monthly locations of the Intertropical Convergence Zone (ITCZ). The emissions of CH_3CCl_3 and sulphur hexafluoride (SF_6) are taken from the transport model intercomparison (TransCom)-CH_4 experiment[6,24,25] and extrapolated for the years after 2008 (Extended Data Fig. 2 and Extended Data Table 1). We show below that the simulated NH–SH gradient of CH_3CCl_3 is equally sensitive to the NH/SH OH ratio in the model with regard to the emissions since 2000. Thus we can use the observed NH–SH CH_3CCl_3 gradient to constrain the NH/SH OH ratio in the model, provided that little uncertainties are contributed by emission estimates and transport parameterization.

Results from both the ACTM_0.99 and ACTM_1.26 simulations reproduce the temporal evolution of CH_3CCl_3 for the period 1994–2011 (Fig. 1), confirming that the global annual mean OH concentrations are in good balance with the 'Control' surface emissions used in the simulations. Concentrations of CH_3CCl_3 have decreased exponentially since the late 1990s as a result of the stringent implementation of the Montreal Protocol and its amendments to mitigate stratospheric ozone depletion[3,12] (Fig. 1a). The simulated CH_3CCl_3 concentration decay rates are not significantly

[1]Department of Environmental Geochemical Cycle Research, JAMSTEC, Yokohama 236 0001, Japan. [2]CAOS, Graduate School of Studies, Tohoku University, Sendai 980 8578, Japan. [3]Wageningen University, Droevendaalsesteeg 3a, 6708 PB, The Netherlands. [4]National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory, Boulder, Colorado 80305, USA. [5]Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA. [6]The Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida 33149, USA. [7]Rutgers, The State University of New Jersey, New Brunswick, New Jersey 08901, USA. [8]National Center for Atmospheric Research (NCAR), Boulder, Colorado 80301, USA. [9]School of Engineering and Applied Science, Harvard University, Cambridge, Massachusetts 02138, USA. [10]Centre for Australian Weather and Climate Research, Commonwealth Scientific and Industrial Research Organisation (CSIRO) Oceans and Atmosphere Flagship, Aspendale, Victoria 3195, Australia. [11]National Institute of Polar Research, 10-3, Midoricho, Tachikawa, Tokyo 190-8518, Japan. [12]CIRES, University of Colorado, Boulder, Colorado 80309, USA. [13]School of Chemistry, University of Bristol, Cantock's Close, BS8 1TS, UK. [14]Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [15]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, USA.

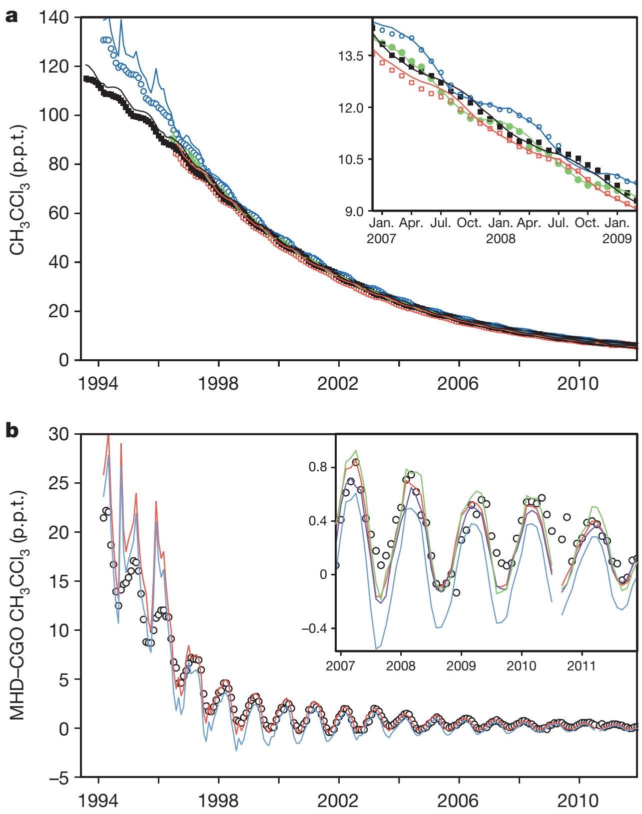


Figure 1 | Temporal evolution of measured (symbols) and simulated (lines) CH_3CCl_3 in the atmosphere. **a**, Monthly mean concentrations at MHD (blue), RPB (green), SMO (red) and CGO (black). Observations (symbols) are taken at four AGAGE sites using GC–MD, and the ACTM simulations (lines) correspond to the 'Control' case of total emissions and annual mean OH. **b**, MHD–CGO concentration differences are shown in comparison with ACTM_0.99 (red) and ACTM_1.26 (blue) simulations. Note that because of the coarse horizontal resolution (T42 spectral truncations; $\sim 2.8^\circ \times 2.8^\circ$) site representation errors are large for Mace Head when intense emissions occurred over Western Europe, for example until 2000 for CH_3CCl_3. ACTM_0.99 simulation at a horizontal resolution of T106 spectral truncations ($\sim 1.1^\circ \times 1.1^\circ$; ACTM_T106 in inset to **b**; green) for the period 2002–2011 shows no significant difference for CH_3CCl_3 from the T42 resolution run, indicating that site representation error does not affect our results. The inset to **b** also shows a model simulation (ACTM_UNEP; purple) using a different emission distribution, based on countries reporting to the United Nations Environment Programme (UNEP), but with identical global emission totals and OH distribution as for ACTM_0.99. The model lines are broken because of a missing observation in August 2010 at CGO. Representative CH_3CCl_3 emission distributions for the 'Control' and 'UNEP' cases are shown in Extended Data Fig. 2. Similar CH_3CCl_3 concentration gradients, based on a greater number of NOAA flask sampling sites, are shown in Extended Data Fig. 3.

different, namely 18.28 ± 0.14% per year (average ± 1σ as interannual variability) for ACTM_0.99 and 17.27 ± 0.13% per year of the annual mean concentrations for ACTM_1.26, compared with the observed 17.85 ± 0.29% per year during 2002–2011 at five Advanced Global Atmospheric Gases Experiment (AGAGE) sites and nine National Oceanic and Atmospheric Administration (NOAA) sites (Extended Data Table 2a). The average CH_3CCl_3 lifetimes are calculated to be 4.91 ± 0.03 years for ACTM_0.99 and 5.19 ± 0.03 years for ACTM_1.26. These lifetimes agree well with the observation-based lifetimes for given inventory emissions of 5.0 (range 4.87–5.23) years[3,26].

Given specified source distribution and magnitude, the meridional CH_3CCl_3 concentration gradients are controlled mainly by loss due to reaction with OH and by meridional transport[23]. This is because the local lifetimes of 1–3 years in the tropical troposphere are of similar magnitude as the interhemispheric transport time of 1.3 years in the ACTMs[6,22]. Figure 1b shows CH_3CCl_3 concentration gradients between Mace Head, Ireland (MHD), and Cape Grim, Australia (CGO). Results from ACTM_0.99 reveal a closer agreement with the observed MHD–CGO CH_3CCl_3 concentration gradients than with those for ACTM_1.26, given the set of 'Control' global emissions and global mean OH concentrations (also noted using NOAA data from multiple sites; see Extended Data Fig. 3). The differences between ACTM_1.26 and ACTM_0.99 are readily apparent for the 2000s, when the yearly emissions of CH_3CCl_3 are less than 3% of the atmospheric burden, compared with the early 1990s, when yearly emissions were as large as ~20% of the burden. Sensitivity simulations are conducted using 'Control' global emissions and mean OH at increased horizontal resolution (ACTM_T106) and with a different spatial distribution of CH_3CCl_3 emissions (ACTM_UNEP). ACTM_T106 and ACTM_UNEP show equally good agreement with observed concentration gradients between MHD and CGO (Fig. 1b, inset), suggesting model resolution and source distribution are, unlike the NH/SH OH ratio, not strong drivers for the NH–SH CH_3CCl_3 concentration gradient. Generally, the ratio of annual mean MHD–CGO CH_3CCl_3 gradients to annual mean concentrations is extremely stable at 2.87 ± 0.41% for AGAGE gas chromatograph–multi detector (GC–MD) data during the 2000s. The ACTM_0.99 simulation well predicted this ratio (2.88 ± 0.19%), whereas it is only 0.54 ± 0.22% for ACTM_1.26. The ratio of the MHD–CGO gradient to the mean CH_3CCl_3 concentration decreased rapidly in the 1990s, from ~34% in 1990 to ~2.9% in 1999 and the 2000s, in proportion to the decrease in global total emissions, mostly occurring in the NH.

Because no formal emission inventory of CH_3CCl_3 exists, uncertainties in the NH/SH emission ratio and in the global totals should be quantified. We find that the MHD–CGO CH_3CCl_3 differences are not particularly sensitive to NH/SH emission ratios of 10 or greater (~16.6 in the 'Control' case) in the period 2004–2011. To assess the uncertainties contributed by the global mean OH concentration and total CH_3CCl_3 emission jointly, we derive a linear relationship between the two (percentage lifetime change = −3.9 × percentage emission change) for simulating the observed CH_3CCl_3 growth rate (Extended Data Fig. 4). We find that ACTM_0.99 produces a minimum for the model-observation mismatch (J) for the magnitude of global emission and chemical loss given as the 'Control' case (Fig. 2). A slightly larger mismatch minimum is observed for ACTM_1.26 when we consider +20% chemical loss and +78% global CH_3CCl_3 emissions during the period 2004–2011 (Fig. 2a, b). However, a significant increase in emission and loss deteriorates the agreement between simulated and observed seasonal cycle amplitudes in the CH_3CCl_3 concentration difference between MHD–CGO or Alert, Canada (ALT), and Palmer Station, Antarctica (PSA) (Fig. 2c). Thus significantly larger global CH_3CCl_3 emissions than considered in the 'Control' case can be ruled out for the period 2004–2011, as opposed to the period of the late 1990s (ref. 8), and therefore the possibility of significantly more OH in the NH than in the SH is also deemed inconsistent with these observations and their simulation.

To exclude erroneous interhemispheric transport in the ACTM model, we briefly present results for SF_6, which is purely anthropogenic and chemically inert in the troposphere, and thus comprises an excellent tracer of atmospheric transport given its relatively high emission rate relative to the global burden and well-known meridional emission distribution[22,27,28]. The meridional transport of SF_6 in ACTM has been validated extensively using surface sites[6,22]. Further, fine-grained measurements of various species during High-performance Instrumented Airborne Platform for Environmental Research (HIAPER) Pole-to-Pole Observation (HIPPO) campaigns[13] represent zonal mean cross-sections of the troposphere (Extended Data Fig. 5). In general, ACTM-simulated SF_6 matches the observations well along all flight paths: the simulated SF_6 values frequently lie within the variability observed by the three instruments that recorded SF_6 on the HIPPO aircraft (Extended Data Fig. 6). The ACTM simulated NH–SH differences agree within ±0.02 parts per trillion (p.p.t.) or ±6% of the observed values for each HIPPO campaign, averaged over all latitudes and 1–3 km altitude for the three instruments (Extended Data Table 3).

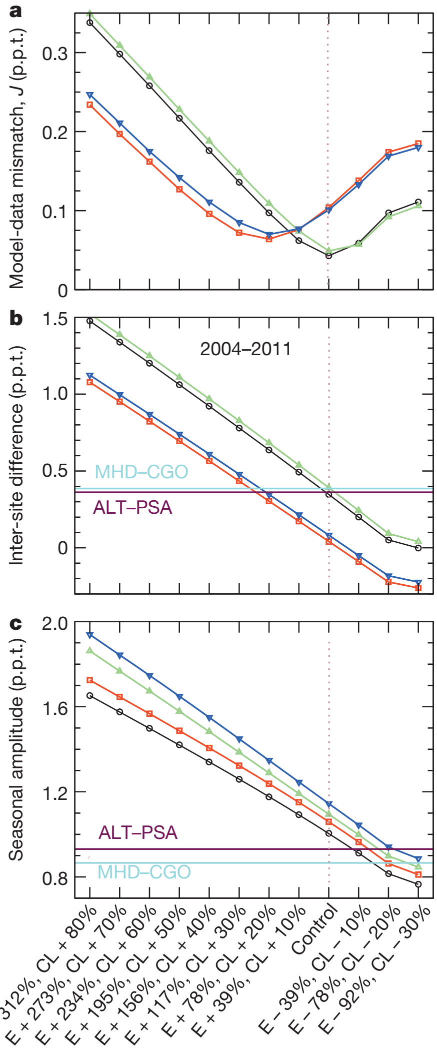


Figure 2 | Role of global total emissions and chemical loss on inter-site CH_3CCl_3 differences. **a**, Model–measurement mismatch (derived as $J = \sqrt{\{[(C_N - C_S)_{model} - (C_N - C_S)_{measured}]^2/N\}}$; C_N and C_S are CH_3CCl_3 concentrations in the NH and SH, respectively, and N is the number of data points) as a function of total emissions (E) and chemical loss (CL) due to global mean OH abundance varied together in a manner consistent with the observed global decline in CH_3CCl_3 concentrations. The mismatch is shown in terms of standard deviations of simulated ALT–PSA (black, ACTM_0.99; red, ACTM_1.26) and MHD–CGO (green, ACTM_0.99; blue, ACTM_1.26) CH_3CCl_3 concentration differences with respect to measurements as monthly averages over the period 2004–2011. **b**, **c**, Annual means of inter-site difference at monthly intervals (**b**) and peak-to-trough seasonal cycle amplitude in the inter-site difference (**c**), to decompose the contribution of E and CL to the model–measurement mismatch. The observed values are shown by horizontal purple lines for ALT–PSA and light blue lines for MHD-CGO. These results are independent of sampling network (MHD–CGO from AGAGE GC–MD and ALT–PSA from NOAA). All the sensitivity simulations were for 2001–2011. Simulations for 2001–2003 have been considered as spin-up, and are excluded for calculating statistics.

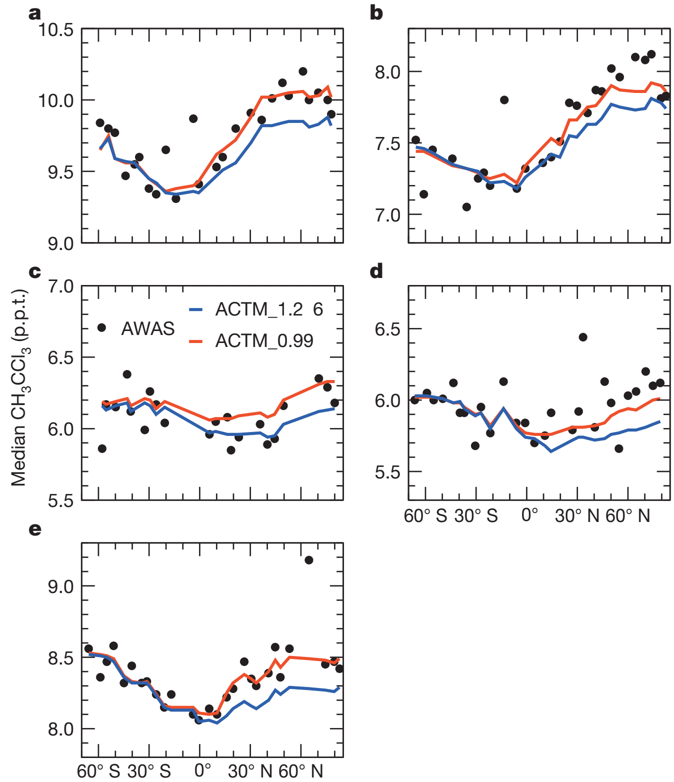


Figure 3 | Meridional gradients of CH_3CCl_3 during five HIPPO campaigns suggest that the NH/SH OH ratio is close to 1. Latitudinal distributions of CH_3CCl_3 are shown as measured from the Advanced Whole Air Sampling (AWAS) flask air (black), and as simulated by ACTM_0.99 (red) and ACTM_1.26 (blue) with 'Control' global emissions and global mean OH concentrations. **a**, HIPPO 1, 12–23 January 2009; **b**, HIPPO 3, 26 March to 15 April 2010; **c**, HIPPO 4, 16 June to 10 July 2011; **d**, HIPPO 5, 19 August to 8 September 2011; **e**, HIPPO 2, 2–21 November 2009. The panels are arranged in seasonal order. The median concentrations are shown at 5° latitude intervals for a 1–4-km altitude range for the meridional gradients. The *y*-axis range is maintained at 1.5 p.p.t., however, the absolute values differ, reflecting time differences between the campaigns. Both ACTM results are adjusted to the mean observed values corresponding to >25° S and the altitude range 1–4 km for each of the HIPPO campaigns separately (+0.07, +0.05, −0.24, +0.05 and −0.05 for ACTM_0.99, and −0.90, −0.85, −1.09, −0.75 and −0.80 for ACTM_1.26 for HIPPO 1–5, respectively), to allow for uncertainties in decadal emissions and lifetimes of CH_3CCl_3 (and bias in concentration gradients for ACTM_1.26), but this systematic shift with the SH reference does not affect the meridional gradient northward of 25° S.

The effect of the NH/SH OH ratio on NH–SH CH_3CCl_3 concentration gradients is well supported by the HIPPO measurements over the Pacific Ocean. The ACTM_0.99 simulated CH_3CCl_3 meridional gradients, given the 'Control' emissions (NH–SH = 0.19 p.p.t. averaged over all HIPPO campaigns), are in close agreement with the HIPPO observations (0.21 p.p.t.), with model-observation differences mostly <0.2 p.p.t. (or <2%) at all latitudes, during all seasons (Fig. 3). However, the results using ACTM_1.26 show systematically smaller NH–SH CH_3CCl_3 differences (0.05 p.p.t. averaged over all HIPPO campaigns), suggesting that the loss due to the CH_3CCl_3 + OH reaction is too high in the NH troposphere or too low in the SH. The poleward increase in CH_3CCl_3 from the SH subtropics is caused mainly by the greater abundance of OH in the SH tropics than in the subtropics in austral summer (HIPPO 1, 2 and 3).

For estimating the possible range of the annual mean NH/SH OH ratio, we performed nine sensitivity simulations using synthetic OH distributions (Extended Data Table 2b) for the period 2001–2011. The synthetic OH distributions are prepared by adding and subtracting sine functions with a phasing of 2 × latitude for ACTM_0.99 and mixing the two OH distributions. Figure 4 shows the dependence of average CH_3CCl_3 concentration gradients between NH and SH on the NH/SH OH ratio. Using the linear fits (shown as lines in Fig. 4) to the model simulations (open symbols), annual NH/SH OH ratios are calculated on the basis of the observed CH_3CCl_3 concentration gradients (filled symbols) averaged over the period 2004–2011 (2009–2011 for HIPPO). The average ($\pm 1\sigma$ of annual values) of NH/SH OH ratios are estimated to be 0.98 ± 0.12, 0.97 ± 0.11 and 0.96 ± 0.63 using the CH_3CCl_3 concentration gradients between MHD and CGO from AGAGE (average of GC–MD and Medusa instruments), ALT and PSA from NOAA flask samples, and NH and

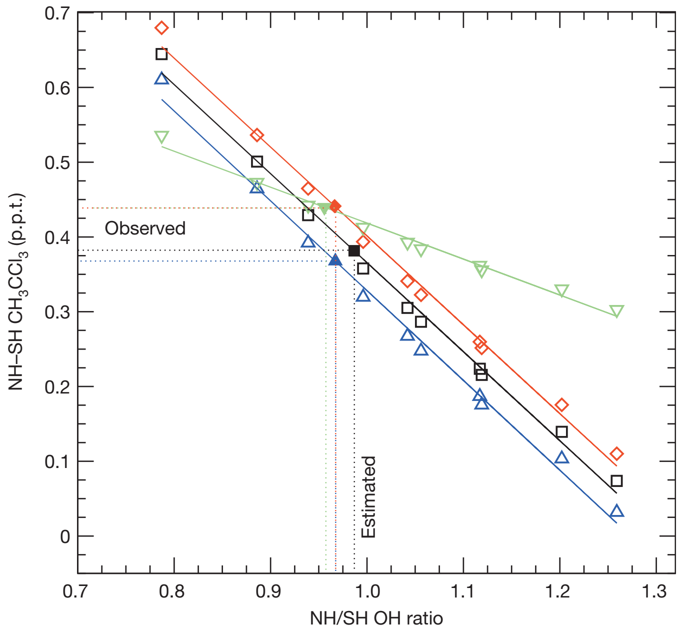


Figure 4 | Estimation of the NH/SH OH concentration ratio from CH_3CCl_3 interhemispheric gradients. NH–SH CH_3CCl_3 concentration differences for different measurement data sets (black, AGAGE GC-MD, MHD–CGO; red, AGAGE Medusa, MHD–CGO; blue, NOAA flask, ALT–PSA; green, HIPPO, between 30° N and 30° S) based on the ACTM sensitivity simulations for various NH/SH OH ratios during the period 2004–2011 considering the case of 'Control' global total emissions and global mean OH concentrations. ACTM simulation results (open symbols) using different OH distributions (ACTM_0.99; ACTM_0.99 ± sine functions; ACTM_0.99 and ACTM_1.26 mixtures; and ACTM_1.26; Extended Data Table 2b) are sampled for the AGAGE GC–MD, AGAGE Medusa, NOAA flasks and HIPPO sampling locations. The observed NH–SH CH_3CCl_3 concentration gradients (closed symbols) are calculated using MHD and CGO for AGAGE, ALT and PSA for NOAA flasks, and averages of data in the latitudes polewards of 30° in the altitude range 1–4 km for HIPPO, which are then used for calculating the NH/SH OH ratio with the fitted lines (GC–MD, $y = 1.556 - 1.191x$; Medusa, $y = 1.589 - 1.188x$; flask, $y = 1.589 - 1.188x$; HIPPO, $y = 0.899 - 0.481x$). Model outputs for the Medusa and GC–MD measurements differ slightly because gaps in the records from the two instruments during the 2004–2011 time period are not coincident.

SH averages (1–4 km altitude, latitudes >30°) from HIPPO, respectively. Combining the results from the AGAGE and NOAA surface network, the decadal average NH/SH OH ratio is 0.97 ± 0.12. The decadal average NH/SH OH ratio estimated from the surface network is in excellent agreement with that estimated from five HIPPO campaigns covering greater geographical areas and vertical extents of both hemispheres but sampling over a briefer period. The uncertainty of about ±13% for the surface networks includes the measurement and model errors (<3%), emission uncertainties, and interannual and seasonal variations in OH within each of the hemispheres, although relative contributions of emission uncertainty and OH variations cannot be quantified.

The precise and well-calibrated measurements from different networks, combined with a transport model that accurately describes interhemispheric transport and modelled emissions, allow us to conclude that a global OH distribution with substantially more OH in the NH is inconsistent with CH_3CCl_3 observations. Our result of an NH/SH OH ratio of close to 1 is in strong contrast to higher modelled OH in the NH[4]. Our results may be explained in various ways. Either NH OH sources are overestimated, possibly owing to O_3 that is biased high in the NH[4], or NH OH sinks are underestimated[29]. Alternatively, SH OH sources may be underestimated[30] or SH OH sinks are overestimated (less likely). Further refinements of OH distributions and OH budgets are required for an accurate estimation of surface emissions of many important short-lived species that affect the Earth's radiative budget and air pollution chemistry. For example, to match the observed interhemispheric gradient in atmospheric CH_4 (based on data from the TransCom experiment[6]), CH_4 emissions in the NH have to be increased from 398 to 430 Tg of CH_4 per year, and decreased from 151 to 119 Tg of CH_4 per year in the SH if ACTM_1.26 OH is used instead of ACTM_0.99. Our results also imply that top-down emission estimates of reactive species (for example, CO, NO_x and SO_x) in key emitting countries in the NH are probably overestimated if OH fields are used with an NH/SH OH ratio much larger than 1.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 21 February; accepted 28 July 2014.

1. Levy, H. Normal atmosphere: large radical and formaldehyde concentrations predicted. *Science* **173,** 141–143 (1971).
2. Crutzen, P. J. in *Tropospheric Ozone: Regional and Global Scale Interactions* (ed. Isaksen, I. S. A.) 3–11 (Reidel, 1988).
3. World Meteorological Organization. *Scientific Assessment of Ozone Depletion: 2010.* (Global Ozone Research and Monitoring Project, Report no. 52, 2011).
4. Naik, V. *et al.* Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP). *Atmos. Chem. Phys.* **13,** 5277–5298 (2013).
5. Miyazaki, K. *et al.* Global NO_x emission estimates derived from an assimilation of OMI tropospheric NO_2 columns. *Atmos. Chem. Phys.* **12,** 2263–2288 (2012).
6. Patra, P. K. *et al.* TransCom model simulations of CH_4 and related species: linking transport, surface flux and chemical loss with CH_4 variability in the troposphere and lower stratosphere. *Atmos. Chem. Phys.* **11,** 12813–12837 (2011).
7. Brenninkmeijer, C. A. M. *et al.* Interhemispheric asymmetry in OH abundance inferred from measurements of atmospheric ^{14}CO. *Nature* **356,** 50–52 (1992).
8. Montzka, S. A. *et al.* New observational constraints for atmospheric hydroxyl on global and hemispheric scales. *Science* **288,** 500–503 (2000).
9. Krol, M. C. & Lelieveld, J. Can the variability in tropospheric OH be deduced from measurements of 1,1,1-trichloroethane (methyl chloroform)? *J. Geophys. Res.* **108,** 4125 (2003).
10. Prinn, R. G. *et al.* Evidence for variability of atmospheric hydroxyl radicals over the past quarter century. *Geophys. Res. Lett.* **32,** L07809 (2005).
11. Prinn, R. G. *et al.* A history of chemically and radiatively important gases in air deduced from ALE/GAGE/AGAGE. *J. Geophys. Res.* **115,** 17751–17792 (2000).
12. Montzka, S. A. *et al.* Small interannual variability of global atmospheric hydroxyl. *Science* **331,** 67–69 (2011).
13. Wofsy, S. C. *et al.* HIAPER Pole-to-Pole Observations (HIPPO): fine grained, global scale measurements for determining rates for transport, surface emissions, and removal of climatologically important atmospheric gases and aerosols. *Phil. Trans. R. Soc. A* **369,** 2073–2086 (2011).
14. Kanaya, Y. *et al.* Chemistry of OH and HO_2 radicals observed at Rishiri Island, Japan, in September 2003: missing daytime sink of HO_2 and positive nighttime correlations with monoterpenes. *J. Geophys. Res.* **112,** D11308 (2007).
15. Lelieveld, J. *et al.* Atmospheric oxidation capacity sustained by a tropical forest. *Nature* **452,** 737–740 (2008).
16. Hofzumahaus, A. *et al.* Amplified trace gas removal in the troposphere. *Science* **324,** 1702–1704 (2009).
17. Elshorbany, Y. F. *et al.* HO_x budgets during HOxComp: A case study of HO_x chemistry under NO_x-limited conditions. *J. Geophys. Res.* **117,** D03307 (2012).
18. Krol, M. C. *et al.* Global OH trend inferred from methyl chloroform measurements. *J. Geophys. Res.* **103,** 10697–10711 (1998).
19. Prinn, R. G. *et al.* Evidence for significant variations of atmospheric hydroxyl radicals in the last two decades. *Science* **292,** 1882–1888 (2001).
20. Krol, M. C. *et al.* What can ^{14}CO measurements tell us about OH? *Atmos. Chem. Phys.* **8,** 5033–5044 (2008).
21. Sudo, K. *et al.* CHASER: a global chemical model of the troposphere. 1. Model description. *J. Geophys. Res.* **107,** 4339 (2002).
22. Patra, P. K. *et al.* Transport mechanisms for synoptic, seasonal and interannual SF_6 variations and 'age' of air in troposphere. *Atmos. Chem. Phys.* **9,** 1209–1225 (2009).
23. Spivakovsky, C. *et al.* Three-dimensional climatological distribution of tropospheric OH: update and evaluation. *J. Geophys. Res.* **105,** 8931–8980 (2000).
24. McCulloch, A. & Midgley, P. M. The history of methyl chloroform emissions: 1951–2000. *Atmos. Environ.* **35,** 5311–5319 (2001).
25. EDGAR4.2. Emission Database for Global Atmospheric Research (EDGAR), release version 4.2. (European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency, 2011).
26. Rigby, M. *et al.* Re-evaluation of the lifetimes of the major CFCs and CH_3CCl_3 using atmospheric trends. *Atmos. Chem. Phys.* **13,** 2691–2702 (2013).
27. Maiss, M. *et al.* Sulfur hexafluoride—a powerful new atmospheric tracer. *Atmos. Environ.* **30,** 1621–1629 (1996).
28. Waugh, D. W. *et al.* Tropospheric SF_6: age of air from the northern hemisphere mid-latitude surface. *J. Geophys. Res.* **118,** 11429–11441 (2013).
29. Mao, J. *et al.* Radical loss in the atmosphere from Cu–Fe redox coupling in aerosols. *Atmos. Chem. Phys.* **13,** 509–519 (2013).

30. Taraborrelli, D. *et al.* Hydroxyl radical buffered by isoprene oxidation over tropical forests. *Nature Geosci.* **5,** 190–193 (2012).

Acknowledgements We thank the HIPPO science team and the crew and support staff at the NCAR Research Aviation Facility, and all the laboratory staff working for AGAGE and NOAA measurement networks. This work is partly supported by the Japan Society for the Promotion of Science/Grants-in-Aid for Scientific Research (KAKENHI) Kiban-A (grant no. 22241008) and Ministry of Education, Culture, Sports, Science and Technology (MEXT) Arctic GRENE projects. NCAR is sponsored by the National Science Foundation (NSF). HIPPO was supported by NSF grants ATM-0628575, ATM-0628519, ATM-0628388 ATM-0628452 and ATM-1036399, by NASA award NNX11AF36G, and by NCAR. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF, NOAA or NASA. M.C.K. is supported by EU FP7 project PEGASOS. AGAGE is supported principally by NASA grants to Massachusetts Institute of Technology (NNX11AF17G) and Scripps Institution of Oceanography (NNX11AF16G) and also by NOAA and the CSIRO. Mace Head is supported by the Department of Energy and Climate Change, award GA0201. We thank the CSIRO Oceans and Atmosphere Flagship and the Bureau of Meteorology for Cape Grim project funding. NOAA flask measurements are supported in part by NOAA's Climate Program Office and its Atmospheric, Chemistry, Carbon Cycle and Climate Program.

Author Contributions P.K.P., M.K., S.A.M., B.X., B.B.S., B.R.L., T.A. and A.G. designed the model experiments and performed data analysis. T.A., E.L.A., S.A.M., B.B.S., J.W.E., P.J.F., E.J.H., D.F.H., P.B.K., B.R.M., F.L.M., J.M., S.O.D., R.G.P., L.P.S., H.J.W., R.F.W., S.C.W. and D.Y. conducted measurements. All co-authors participated in writing the manuscript and contributed through discussions.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.K.P. (prabir@jamstec.go.jp).

LETTER

doi:10.1038/nature13546

A major advance of tropical Andean glaciers during the Antarctic cold reversal

V. Jomelli[1], V. Favier[2], M. Vuille[3], R. Braucher[4], L. Martin[5], P.-H. Blard[5], C. Colose[3], D. Brunstein[1], F. He[6], M. Khodri[7], D. L. Bourlès[4], L. Leanni[4], V. Rinterknecht[8], D. Grancher[1], B. Francou[9], J. L. Ceballos[10], H. Fonseca[11], Z. Liu[12] & B. L. Otto-Bliesner[13]

The Younger Dryas stadial, a cold event spanning 12,800 to 11,500 years ago, during the last deglaciation, is thought to coincide with the last major glacial re-advance in the tropical Andes[1]. This interpretation relies mainly on cosmic-ray exposure dating of glacial deposits. Recent studies, however, have established new production rates[2–4] for cosmogenic ^{10}Be and ^{3}He, which make it necessary to update all chronologies in this region[1,5–15] and revise our understanding of cryospheric responses to climate variability. Here we present a new ^{10}Be moraine chronology in Colombia showing that glaciers in the northern tropical Andes expanded to a larger extent during the Antarctic cold reversal (14,500 to 12,900 years ago) than during the Younger Dryas. On the basis of a homogenized chronology of all ^{10}Be and ^{3}He moraine ages across the tropical Andes, we show that this behaviour was common to the northern and southern tropical Andes. Transient simulations with a coupled global climate model suggest that the common glacier behaviour was the result of Atlantic meridional overturning circulation variability superimposed on a deglacial increase in the atmospheric carbon dioxide concentration. During the Antarctic cold reversal, glaciers advanced primarily in response to cold sea surface temperatures over much of the Southern Hemisphere. During the Younger Dryas, however, northern tropical Andes glaciers retreated owing to abrupt regional warming in response to reduced precipitation and land–surface feedbacks triggered by a weakened Atlantic meridional overturning circulation. Conversely, glacier retreat during the Younger Dryas in the southern tropical Andes occurred as a result of progressive warming, probably influenced by an increase in atmospheric carbon dioxide. Considered with evidence from mid-latitude Andean glaciers[16], our results argue for a common glacier response to cold conditions in the Antarctic cold reversal exceeding that of the Younger Dryas.

The general warming trend during deglaciation was interrupted by cooler conditions in the Southern Hemisphere during the Atlantic cold reversal (ACR). Conversely, temperature records from Greenland reveal warm conditions during the ACR (termed the Bølling–Allerød interstadial in the Northern Hemisphere), followed by the cold Younger Dryas event. The response of tropical Andean glaciers to these rapid and nonlinear climate changes remains puzzling. A recent review of published data[1] suggests that tropical Andean glaciers recorded a Younger Dryas signal, a view supported by several ^{10}Be chronologies[5,8,13,15]. However, the dating accuracy of these glacier fluctuations is questionable because ^{10}Be chronologies are affected by large uncertainties (>10%) associated with the cosmogenic production rates. This prevents unambiguous attributions of glacier response to the ACR and Younger Dryas events. Indeed, at least three scaling schemes using different sea-level, high-latitude ^{10}Be production rates were considered in establishing these chronologies. More importantly, recent calibration studies for the first time established local production rates for cosmogenic ^{3}He and ^{10}Be in the high tropical Andes[2–4]. These new developments imply that all previously published moraine ages need to be reconsidered and that the mechanisms leading to glacial advance during the ACR and Younger Dryas events warrant further investigation.

Here we present a new chronology of eight prominent moraines of the Ritacuba Negro glacier (Colombia, Sierra Nevada del Cocuy) deposited during the 'late glacial', that is, the later stages of the last deglaciation. Forty-six ^{10}Be cosmic-ray exposure (CRE) ages were obtained from boulders collected on the moraines and roches moutonnées (Fig. 1 and Methods). Analytic uncertainties on the entire set of CRE ages averaged $6 \pm 6\%$. The Ritacuba Negro glacier chronology was compared with a recalculated data set comprising 246 published ^{10}Be and 12 ^{3}He ages (Supplementary Information) obtained from 47 moraines[1,5–15] sampled on one glacier in the northern tropical Andes (NTA) and 19 glaciers in the southern tropical Andes (STA) over the last 15 kyr. The recalculated data set was standardized using the recently revised local production rate[2] of 3.95 ± 0.18 atoms g^{-1} yr^{-1} with a time-dependent scaling and a specific Andes atmosphere model (Methods). It is important to stress that the production rate used here was calibrated at locations that are comparable in elevation and latitude ranges to the dated moraines. To assess the impact of the different scaling parameters, we report the ages using four different scaling models (Methods).

When used in combination, the new and published ages allow investigation of the following key questions, at the regional scale of the tropical Andes. (1) When did the maximum glacial extents occur over the last 15 kyr in the NTA and the STA, respectively? (2) Did the tropical Andean glaciers show a synchronous behaviour? (3) What climatic mechanisms were driving the observed glacier fluctuations?

The maximum glacial extent of Ritacuba Negro glacier during the late glacial is indicated by the outer and frontal termination moraine M18, located at 3,975 m above sea level, and dates to 13.9 ± 0.3 ^{10}Be kyr ago ($n = 5$) (Fig. 1; ages expressed in these units are calculated from the measured ^{10}Be concentrations). Upslope, four boulders on moraine M17 are internally consistent and yield a mean CRE age of 14.0 ± 0.3 ^{10}Be kyr. Seven samples collected on the large moraine M16 yield a mean CRE age of 13.4 ± 0.3 ^{10}Be kyr. These three moraines indicate several advances or stillstands during the ACR. Upslope from M16, a very large accumulation is composed of three small moraines: M15 formed 11.8 ± 0.2 ^{10}Be kyr ago ($n = 4$), at the very end of the Younger Dryas, and M14 and M13 yield respective mean CRE ages of 11.3 ± 0.1 ^{10}Be kyr ($n = 9$) and 11.0 ± 0.4 ^{10}Be kyr ($n = 4$). Two samples on a roche moutonnée confirm the chronology with a mean age of 11.1 ± 0.2 ^{10}Be kyr. M12, which is roughly 350 m upslope from M13, dates to 1.2 ± 0.1 ^{10}Be kyr

[1]Université Paris 1 Panthéon-Sorbonne, CNRS Laboratoire de Géographie Physique, 92195 Meudon, France. [2]Université Grenoble Alpes, LGGE, UMR 5183, F-38041 Grenoble, France. [3]Department of Atmospheric and Environmental Sciences, University at Albany, Albany, New York 12222, USA. [4]Aix-Marseille Université, CNRS-IRD-Collège de France, CEREGE UM34, 13545 Aix-en-Provence, France. [5]CNRS, Centre de Recherches Pétrographiques et Géochimiques, UMR 7358, Université de Lorraine, BP 20, Vandoeuvre-lès-Nancy 54501, France. [6]Center for Climatic Research, Nelson Institute for Environmental Studies, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA. [7]IRD-Laboratoire d'Océanographie et du Climat: Expérimentation et Approche numérique, Université Pierre et Marie Curie, F-75252 Paris Cedex 05, France. [8]School of Geography and Geosciences Irvine Building, University of St Andrews, St Andrews KY16 9AL, UK. [9]Institut de Recherche pour le Développement, CP 9214, La Paz, Bolivia. [10]Institute for Hydrology, Meteorology and Environmental Studies, Bogota, 07603, Colombia. [11]Escuela de Ingeniería Geológica, UPTC Sede Seccional Sogamoso, Sogamoso, 152211, Colombia. [12]Center for Climatic Research and Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA. [13]Climate and Global Dynamics Division, National Center for Atmospheric Research, Boulder, Colorado 80305, USA.

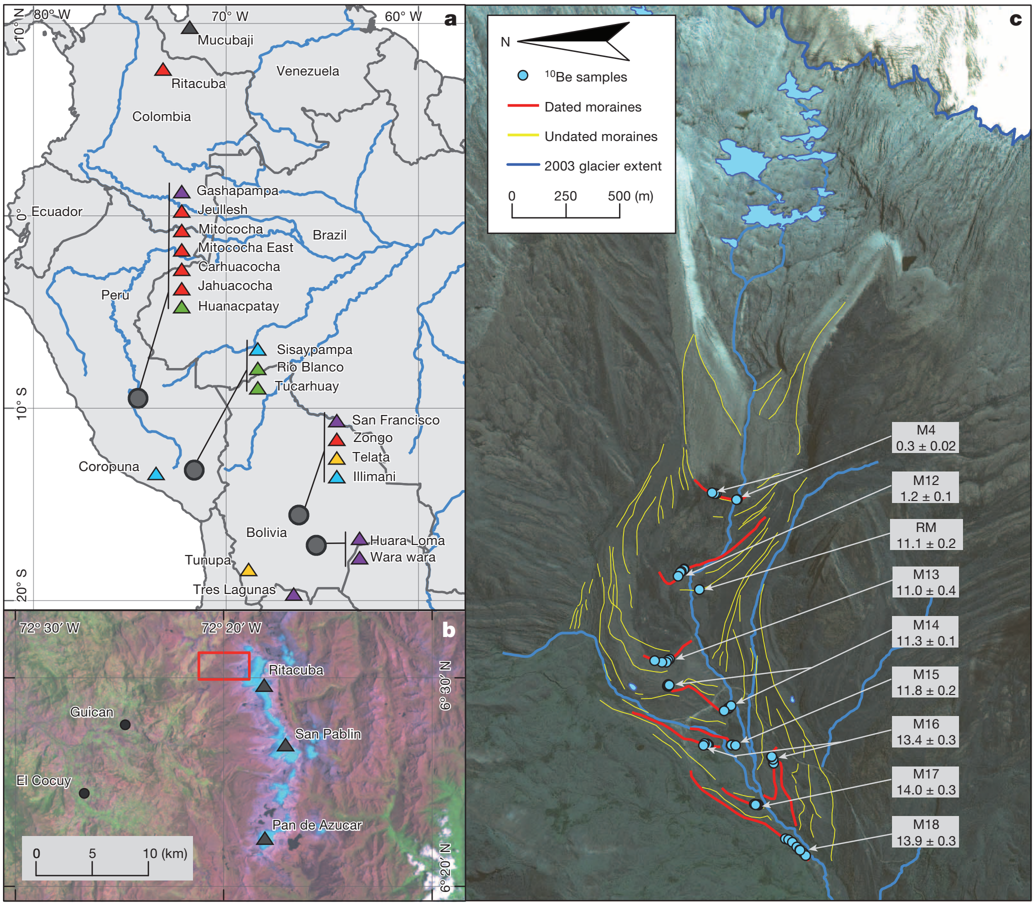


Figure 1 | The Ritacuba Negro glacier and studied sites. **a**, Location of the homogenized ^{10}Be and ^{3}He moraine record sites covering the northern and southern tropical Andes, with the largest glacial advance dated to during the ACR or possibly before (considering uncertainties) in purple; those during the ACR in red; those during the ACR or the Younger Dryas (considering uncertainties) in orange; those during the Younger Dryas in blue; those during the Holocene in green; and rejected chronology in black (Methods). **b**, Location of the northern tropical Ritacuba Negro glacial valley in the Cordillera de Cocuy (red square), with filled triangles indicating summits. **c**, Map of the Ritacuba Negro glacier, showing dated and undated moraines (prefix M indicates a main moraine as discussed in the text and RM means roche moutonnée; units, ^{10}Be kyr; Supplementary Information and Methods), the location of ^{10}Be samples (blue dots), the snout of the Ritacuba Negro glacier in 2003 (thick blue line). The uncertainties associated with the ages account for analytical uncertainties only (1 s.d.).

($n = 2$). The innermost dated moraine of the Ritacuba Negro sequence is located about 2.5 km from the present frontal position of 4,660 m above sea level. Three boulders from this ridge yield a mean CRE age of 264 ± 23 ^{10}Be yr. Finally, three small, fresh moraines were formed during the twentieth century. Among the 46 samples, six were rejected as outliers on the basis of a χ^2 test reflecting cosmogenic nuclide inheritance from previous exposures and post-depositional erosion processes (two from M18, two from M16, one from M12 and one from M4; Methods).

To evaluate the wider implications of the Ritacuba Negro glacier moraine chronology, we first compare it with indirect evidence of glacier fluctuations derived from lake-level fluctuations in Venezuela[17]. The Venezuelan glacier chronology[5] was not considered because of the uncertainties associated with ^{10}Be CRE ages (Methods). ACR advances (or stillstands) are evident in both records (moraine and lake sediments) at ~14.0 kyr ago. Minor advances (or stillstands) at the end of the Younger Dryas and during the early Holocene can also be detected in both records. However, on the basis of high titanium concentrations, ref. 17 identified a major glacial advance between ~12.8 and 12.1 kyr ago in their record. Such a glacial stillstand may have occurred in the Ritacuba Negro valley (Fig. 2), but, if so, it would necessarily have been smaller than both the ACR advances and the ones occurring at the end of the Younger Dryas. Indeed, there is no moraine dated to between 12.8 and 12.1 kyr ago preserved on Ritacuba Negro valley. However, the moraine M15, dated to 11.8 ± 0.2 kyr ago, could correspond to the end of the Younger Dryas.

We then compared the behaviour of the Ritacuba Negro glacier with 16 STA glacier chronologies that cover the ACR/Younger Dryas period (Figs 1 and 2 and Methods). The data show that seven glaciers have formed moraines at least once during the ACR chronozone *sensu stricto* and that seven others contain moraine deposits, whose dates, within the margin of error, overlap with the ACR period (Methods). Moraine formation implies the obliteration of any older moraines deposited by less extensive glaciation upstream, and the ACR advances correspond to the outermost front positions over the last 14.5 kyr in many locations in Peru, Bolivia and northern Argentina. Consequently, the corresponding ACR glacial stillstands are undoubtedly more extensive than those that occurred later during the Younger Dryas. This comparison thus reveals comparable behaviour between the Ritacuba Negro (NTA) and STA glaciers. Glacial advances during the Younger Dryas were recorded in several cordilleras but were generally slightly smaller than those occurring during the ACR. However, larger advances during the Younger Dryas than during the ACR are observed for three glaciers[9,13,15] (five within the limits of dating uncertainty; Fig. 1) and may result from site-specific conditions.

Three glaciers in our data set contain only Holocene moraines (Fig. 1 and Methods), and suggest that early Holocene glacial extents are observed in the Ritacuba valley and at many STA sites. However, it is clear that mid- and late-Holocene stillstands are very rarely observed, probably because these moraines have been erased by Little Ice Age glacial advances[18]. Hence, a coherent retreat from the ACR extent to the present

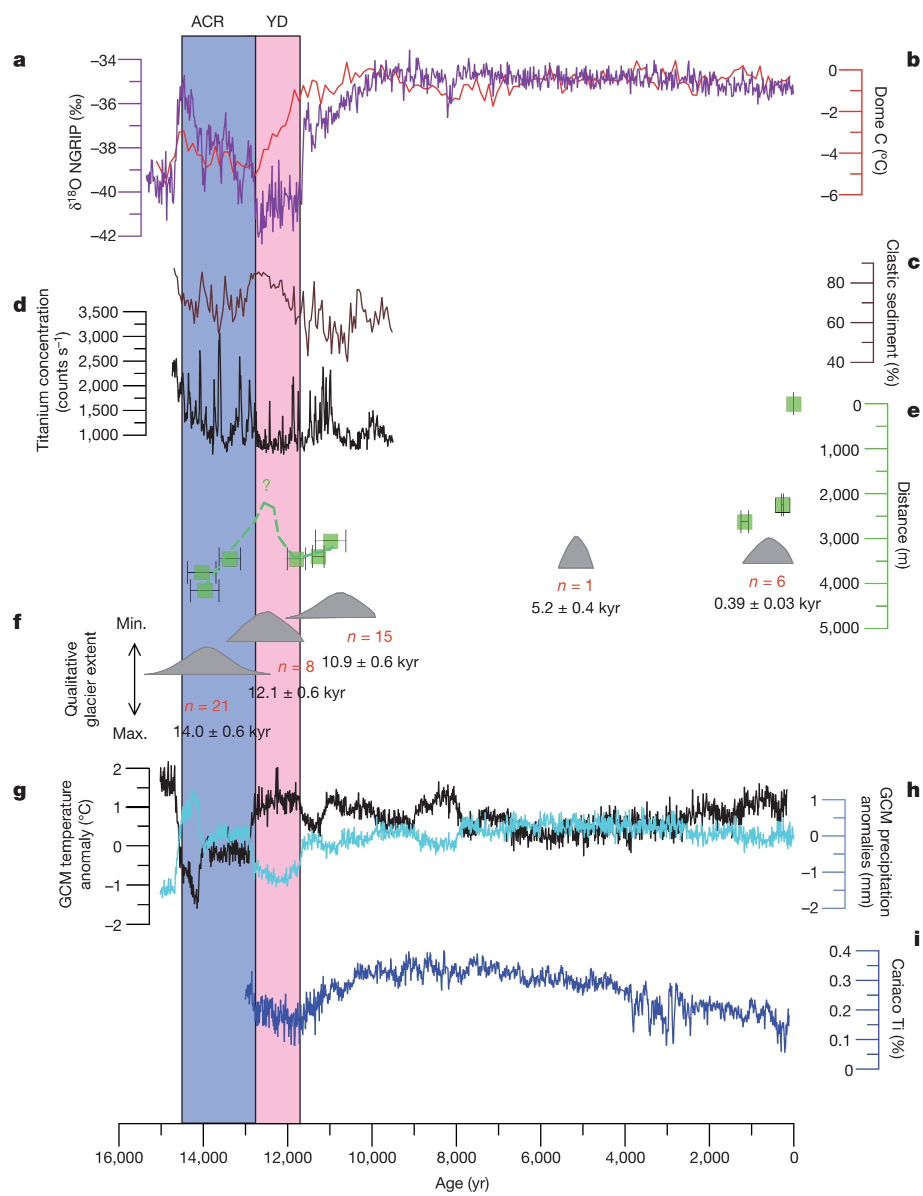


Figure 2 | Changes in the Ritacuba Negro glacier compared with proxy records. **a**, NGRIP $\delta^{18}O$ from ref. 24 (purple line). **b**, Temperature anomalies at EPICA Dome C[24] (red line). **c**, **d**, Clastic sediment (**c**; brown line) and titanium concentration (**d**; black line) from Los Anteojos lake[17] (Venezuela). **e**, Ritacuba Negro glacier front variations relative to extent in 2010 and chronology based on the 40 new ^{10}Be ages documenting the NTA region. Error bars are moraine age uncertainty (1 s.d.). The dashed line shows the possible evolution of the front. **f**, STA moraine ages (based on 246 ^{10}Be surface exposure ages from 19 glaciers; Supplementary Information and Methods). Shaded grey areas correspond to probability distribution functions of moraine ages. Their position on the *y* axis illustrates the progressive general retreat of the glaciers over time. The number of moraines <15 kyr old is shown in red, the mean age of each distribution (uncertainty, 2 s.d.) is shown in black. **g**, CCSM3 temperature anomalies in the Ritacuba region (77–69° W, 2–10° N; black line). **h**, CCSM3 precipitation anomalies in the Ritacuba region (77–69° W, 2–10° N; blue line). Anomalies are with respect to the 13.9 kyr period in the all-forcings run. **i**, Titanium concentration in Cariaco basin sediments[19]. YD, Younger Dryas; ACR, Antarctic cold reversal.

position, interrupted by minor stillstands or re-advances during the Younger Dryas and early Holocene epoch, is observed across the tropical Andes. Together these fluctuations reveal a common trend in glacier size evolution.

The glacier size evolution across the tropics during the ACR/Younger Dryas period is in step with other Southern Hemisphere glaciers such as those in Patagonia and New Zealand[16], and strongly suggests that they mostly result from a common climate driver. The fact that the NTA and STA glacier systems, each exposed to different precipitation regimes[19], display a common evolution suggests that increased temperature served as a dominant control for glacier retreat during the ACR/Younger Dryas period (Figs 2 and 3). This temperature sensitivity is consistent with modern observations which show that temperature affects glacier melt rates through a change in the rain–snow line and albedo feedbacks[20,21].

To explore possible mechanisms responsible for this tropical Andean glacier evolution during the ACR/Younger Dryas period, we analysed the transient simulation of the last deglaciation with the coupled global climate model[22] (GCM) Community Climate System Model version 3 (CCSM3) (Methods). Two studies[23,24] demonstrate that the GCM simulation successfully represents the antiphased hemispheric temperature response to ocean circulation changes during the last deglaciation. The transient simulation indicates a significant warming over the STA region during the deglaciation, interrupted by a minimum 14.1 kyr ago and a smaller decrease in temperature ~12.1 kyr ago (Fig. 3). The temperature change is in good agreement with moraine records in the STA. In the Ritacuba region, however, temperature changed rapidly between 14.1 and 11 kyr ago, with two cold episodes during the ACR and at the end of the Younger Dryas, separated by a warm period during the Younger

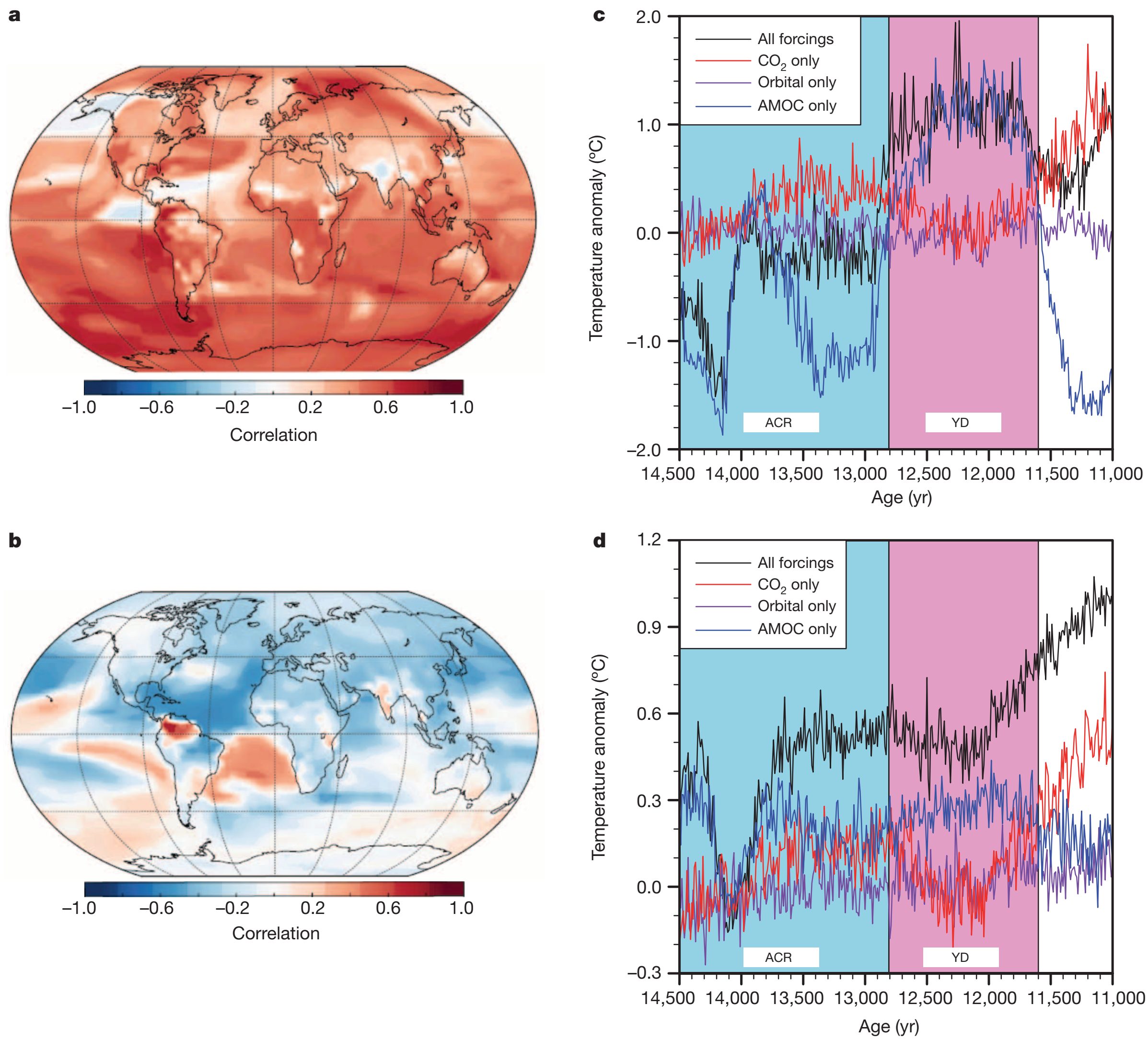


Figure 3 | Decadal temperature variations in the Ritacuba region correlated with global surface temperature. **a**, ACR period. **b**, Younger Dryas period. Statistically insignificant ($P > 0.05$) values are shown in white. **c**, **d**, Temperature evolution simulated with different CCSM3 single-forcing and all-forcing runs in the Ritacuba region (77–69° W, 2–10° N; **c**) and, documenting the STA region, in the Andes of Bolivia[14] (73–65° W, 14–22° S; **d**). The keys show the various forcings used. Anomalies are with respect to the decadal averages at 14.0 kyr in all-forcing runs.

Dryas (Figs 2 and 3). Again this temperature evolution is in agreement with our direct observations of glacier change in the Ritacuba region, but is inconsistent with results from ref. 17, where a cold episode is identified during the main Younger Dryas period on the basis of clastic sediments and pollen collected in a Venezuelan lake. This discrepancy may result from uncertainties in regional GCM simulations or from distinct sensitivities of the different proxies to climate forcing.

To further explore such a hypothesis and better understand this common glacier behaviour, decadal temperature and precipitation variations were correlated with global surface temperature and precipitation fields, respectively. During the ACR, a positive relationship is observed between temperature fluctuations in the Ritacuba region and temperatures over large parts of the Southern Hemisphere. Correlations are most significant at southern high latitudes and in the eastern equatorial Pacific (Fig. 3), with cold sea surface temperatures in the eastern equatorial Pacific being associated with glacier advance, in agreement with present-day observations[25]. During the Younger Dryas, the slowdown of the Atlantic meridional overturning circulation (AMOC) that maintained cold sea surface temperatures in the northern tropical Atlantic produced a very different pattern. In the STA, temperature evolved gradually and in step with the large-scale temperature signal. In the Ritacuba region, however, the continental temperature warmed when the cooling in the Northern Hemisphere occurred during the Younger Dryas (Supplementary Discussion). In the transient simulation, the Younger Dryas warming in the Ritacuba region results from decreased latent heat loss due to reduction of tropical forests, which is caused by the southward shift of the intertropical convergence zone[26] associated with the slowdown of the AMOC during the Younger Dryas (Supplementary Discussion and Extended Data Figs 4–9). Therefore, the temperature increase in the region during the Younger Dryas (Fig. 2) may at least in part be caused by decreased cloudiness and related local land–surface feedbacks such as reduced soil moisture and less evaporative cooling[27,28], as a result of the large-scale reorganization of precipitation over tropical South America. Other mechanisms, such as upwelling Antarctic intermediate water in the eastern tropical Pacific[29] may also have a role in NTA warming during the Younger Dryas. CCSM3 results from simulations[23] that isolate individual forcing components indicate that tropical glacier fluctuations during the ACR/Younger Dryas period were primarily driven by a CO_2 increase superimposed on AMOC variability. AMOC variability was responsible for the abrupt regional climate change observed in the NTA during the Younger Dryas period, whereas temperature changes in the STA carry a predominantly CO_2-forced fingerprint (Fig. 3).

Our results clearly demonstrate that tropical Andean glaciers were impacted by the ACR, consistent with results from recent studies in southern mid latitudes[16], suggesting a common temperature response to this event along the entire Andean cordillera. Regardless of our interpretation, any proposed mechanisms for drivers of deglacial climate change in the tropical Andes must account for the widespread stability of glacier ice during the ACR. Our analyses suggest AMOC variability superimposed on CO_2 forcing as the main drivers of the late deglaciation in the Andes. Finally, our results based on new cosmogenic production rates[2–4] illustrate that most previous chronologies and climate interpretations from tropical glaciers since the LGM may need to be revisited.

METHODS SUMMARY

To compare chronologies from the northern and southern tropical Andes, we homogenized existing late-glacial cosmogenic ^{10}Be and ^{3}He ages younger than 21 kyr (Supplementary Table 1). Beryllium-10 concentrations were normalized against an assigned value of the NIST ^{10}Be/^{9}Be ratio (2.79×10^{-11}). Existing ^{10}Be ages were recalculated using the recent Altiplano production rate of 3.95 ± 0.11 atoms $g^{-1}\,yr^{-1}$. To test the impact of the different parameters involved in the production scaling, we recalculated the age of each moraine from 20 glaciers selected in this study using four models (Supplementary Information). We excluded all glaciers that did not document the ACR–Holocene period. In this process, moraine identification was strictly the same as those documented in cited studies. We used and excluded the same samples as in the cited studies.

We did not use a χ^2 analysis to compare the ages of the moraines because such a test was not used in most previous papers. Instead, we conducted two distinct analyses on the 20 glaciers. The first was done to assess the number of glaciers with the maximum extent belonging to the ACR chronozone, assuming that the youngest moraine (as dated) since the ACR period corresponds to the maximum extent. In this case, we distinguished five different chronozones: pre-ACR, ACR, ACR/Younger Dryas, Younger Dryas and post-Younger Dryas. Each glacier was classified in a single chronozone according to the age of the maximum extent moraine and its uncertainty. The second analysis was conducted on the moraine ages to get the distribution function with time. This time we distinguished three groups of moraines, corresponding to the ACR, Younger Dryas and Holocene chronozones, respectively. Each moraine was classified in one or two distinct chronozones according to age and the associated uncertainties.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 November 2013; accepted 28 May 2014.
Published online 24 August 2014.

1. Rodbell, D. T., Smith, J. A. & Mark, B. G. Glaciation in the Andes during the late glacial and Holocene. *Quat. Sci. Rev.* **28,** 2165–2212 (2009).
2. Blard, P. H., Braucher, R., Lavé, J. & Bourlès, D. Cosmogenic ^{10}Be production rate calibrated against 3He in the high tropical Andes (3800–4900 m, 20–22° S). *Earth Planet. Sci. Lett.* **382,** 140–149 (2013).
3. Blard, P.-H. *et al.* Cosmogenic ^{3}He production rate in the high tropical Andes (3800 m, 20°S): implications for the local last glacial maximum. *Earth Planet. Sci. Lett.* **377–378,** 260–275 (2013).
4. Kelly, M. *et al.* A locally calibrated, late glacial ^{10}Be production rate from a low latitude, high-altitude site in the Peruvian Andes. *Quat. Geochronol*. (in the press).
5. Carcaillet, J., Angel, I., Carrillo, E., Audemard, F. A. & Beck, C. Timing of the last deglaciation in the Sierra Nevada of the Merida Andes, Venezuela. *Quat. Res.* **80,** 482–494 (2013).
6. Smith, J. A., Seltzer, G. O., Farber, D. L., Rodbell, D. T. & Finkel, R. C. Early local last glacial maximum in the tropical Andes. *Science* **308,** 678–681 (2005).
7. Zech, R., Kull, Ch., Kubik, P. W. & Veit, H. LGM and Late Glacial glacier advances in the Cordillera Real and Cochabamba (Bolivia) deduced from ^{10}Be surface exposure dating. *Clim. Past* **3,** 623–635 (2007).
8. Glasser, N. F., Clemmens, S., Schnabel, C., Fenton, C. R. & McHargue, L. Tropical glacier fluctuations in the Cordillera Blanca, Peru between 12.5 and 7.6 ka from cosmogenic ^{10}Be dating. *Quat. Sci. Rev.* **28,** 3448–3458 (2009).
9. Licciardi, J. M., Schaefer, J. M., Taggart, J. R. & Lund, D. C. Holocene glacier fluctuations in the Peruvian Andes indicate northern climate linkages. *Science* **325,** 1677–1679 (2009).
10. Hall, S. R. *et al.* Geochronology of Quaternary glaciations from the tropical Cordillera Huayhuash, Peru. *Quat. Sci. Rev.* **28,** 2991–3009 (2009).
11. Zech, J., Zech, R., Kubik, P. W. & Veit, H. Glacier and climate reconstruction at Tres Lagunas, NW Argentina, based on ^{10}Be surface exposure dating and lake sediment analyses. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **284,** 180–190 (2009).
12. Smith, J. A. & Rodbell, D. T. Cross-cutting moraines reveal evidence for North Atlantic influence on glaciers in the tropical Andes. *J. Quat. Sci.* **25,** 243–248 (2010).
13. Smith, C. A., Lowell, T. V., Owen, L. A. & Caffee, M. W. Late Quaternary glacial chronology on Nevado Illimani, Bolivia, and the implications for paleoclimatic reconstructions across the Andes. *Quat. Res.* **75,** 1–10 (2011).
14. Jomelli, V. *et al.* Irregular tropical glacier retreat over the Holocene epoch driven by progressive warming. *Nature* **474,** 196–199 (2011).
15. Bromley, G. *et al.* Glacier fluctuations in the southern Peruvian Andes during the late-glacial period, constrained with cosmogenic ^{3}He. *J. Quat. Sci.* **26,** 37–43 (2011).
16. Putnam, A. E. *et al.* Glacier advance in southern middle-latitudes during the Antarctic Cold Reversal. *Nature Geosci.* **3,** 700–704 (2010).
17. Stansell, N. D. *et al.* Abrupt Younger Dryas cooling in the northern tropics recorded in lake sediments from the Venezuelan Andes. *Earth Planet. Sci. Lett.* **293,** 154–163 (2010).
18. Jomelli, V. *et al.* Fluctuations of Andean tropical glaciers since the last millennium and palaeoclimatic implications: a review. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **281,** 269–282 (2009).
19. Haug, G. H., Hughen, K. A., Sigman, D. M., Peterson, L. C. & Röhl, U. Southward migration of the Intertropical Convergence Zone through the Holocene. *Science* **293,** 1304–1308 (2001).
20. Favier, V., Wagnon, P. & Ribstein, P. Glaciers of the outer and inner tropics: a different behaviour but a common response to climatic forcing. *Geophys. Res. Lett.* **31,** L16403 (2004).
21. Favier, V., Wagnon, P., Chazarin, J.-P., Maisincho, L. & Coudrain, A. One-year measurements of surface heat budget on the ablation zone of Antizana Glacier 15, Ecuadorian Andes. *J. Geophys. Res.* **109,** D18105 (2004).
22. Liu, Z. *et al.* Transient simulation of deglacial climate evolution with a new mechanism for Bolling-Allerod warming. *Science* **325,** 310–314 (2009).
23. He, F. *et al.* Northern Hemisphere forcing of Southern Hemisphere climate during the last deglaciation. *Nature* **494,** 81–85 (2013).
24. Shakun, J. D. *et al.* Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation. *Nature* **484,** 49–54 (2012).
25. Francou, B., Vuille, M., Favier, V. & Caceres, B. New evidence for an ENSO impact on low-latitude glaciers: Antizana 15, Andes of Ecuador, 0–28°S. *J. Geophys. Res.* **109,** D18106 (2004).
26. Peterson, L. C., Haug, G. H., Hughen, K. A. & Rohl, U. Rapid changes in the hydrologic cycle of the tropical Atlantic during the last glacial. *Science* **290,** 1947–1951 (2000).
27. Leduc, G. *et al.* Moisture transport across central America as a positive feedback on abrupt climatic changes. *Nature* **445,** 908–911 (2007).
28. Davin, E. L. & de Noblet-Ducoudre, N. Climatic impact of global-scale deforestation: radiative versus nonradiative processes. *J. Clim.* **23,** 97–112 (2010).
29. Pierrehumbert, R. T. Climate change and the tropical Pacific: the sleeping dragon wakes. *Proc. Natl Acad. Sci. USA* **97,** 1355–1358 (2000).

Supplementary Information is available in the online version of the paper.

Acknowledgements Financial support was provided by the French ANR El Paso programme no. 10-blan-68-01. The ^{10}Be measurements were performed at the ASTER AMS national facility (CEREGE, Aix en Provence), which is supported by the INSU/CNRS, the French Ministry of Research and Higher Education, IRD and CEA. TRACE21 is supported by the P2C2 programme (NSF), the Abrupt Change Program (DOE), the EaSM programme (DOE) and the INCITE computing programme (DOE and NCAR). We thank M. Arnold, G. Aumaître and K. Keddadouche for their assistance during ^{10}Be measurements.

Author Contributions V.J., D.B., J.L.C. and H.F. conducted the field work on Ritacuba Negro glacier; F.H., Z.L. and B.O.-B. performed the GCM simulations; M.V. and C.C. provided temperature correlation maps; D.L.B., R.B., P.-H.B., L.L. and L.M. participated in producing the cosmogenic data; L.M., P.-H.B. and V.J. updated and homogenized the previously published cosmogenic ages; P.-H.B., V.R., V.J., D.G. and D.L.B. interpreted the cosmogenic ages; and V.J., V.F., M.V., F.H. and D.L.B. contributed to writing the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to V.J. (vincent.jomelli@lgp.cnrs.fr).

doi:10.1038/nature13717

A global strategy for road building

William F. Laurance[1], Gopalasamy Reuben Clements[1,2], Sean Sloan[1], Christine S. O'Connell[3], Nathan D. Mueller[4], Miriam Goosem[1], Oscar Venter[1], David P. Edwards[5], Ben Phalan[6], Andrew Balmford[6], Rodney Van Der Ree[7] & Irene Burgues Arrea[8]

The number and extent of roads will expand dramatically this century[1]. Globally, at least 25 million kilometres of new roads are anticipated by 2050; a 60% increase in the total length of roads over that in 2010. Nine-tenths of all road construction is expected to occur in developing nations[1], including many regions that sustain exceptional biodiversity and vital ecosystem services. Roads penetrating into wilderness or frontier areas are a major proximate driver of habitat loss and fragmentation, wildfires, overhunting and other environmental degradation, often with irreversible impacts on ecosystems[2–5]. Unfortunately, much road proliferation is chaotic or poorly planned[3,4,6], and the rate of expansion is so great that it often overwhelms the capacity of environmental planners and managers[2–7]. Here we present a global scheme for prioritizing road building. This large-scale zoning plan seeks to limit the environmental costs of road expansion while maximizing its benefits for human development, by helping to increase agricultural production, which is an urgent priority given that global food demand could double by mid-century[8,9]. Our analysis identifies areas with high environmental values where future road building should be avoided if possible, areas where strategic road improvements could promote agricultural development with relatively modest environmental costs, and 'conflict areas' where road building could have sizeable benefits for agriculture but with serious environmental damage. Our plan provides a template for proactively zoning and prioritizing roads during the most explosive era of road expansion in human history.

A multitude of factors is promoting rapid road expansion globally, including a quest for valuable resources such as timber, minerals, oil and arable land, and initiatives to increase regional trade, transportation and energy infrastructure[4,7]. Yet, while new roads can promote social and economic development[10,11], they also can open a Pandora's box of environmental problems[2–7]. This is especially the case in pristine or frontier regions, where new roads often dramatically increase land colonization, habitat disruption, and overexploitation of wildlife and natural resources[2–6]. It is broadly understood that the best strategy for maintaining the integrity of wilderness areas is by 'avoiding the first cut'—keeping them road-free[4]—because deforestation is highly contagious spatially[12] and because new roads tend to spawn networks of secondary and tertiary roads that greatly increase the extent of environmental damage[4]. Unfortunately, new roads are now penetrating into many of the world's last surviving wildernesses, including the Amazon[2,5,6,10], New Guinea[13], Siberia[14] and the Congo Basin[3,11,15].

However, some roads generate substantial social and economic benefits with only modest environmental costs. Particularly in developing nations, vast expanses of land have been settled but have low agricultural productivity because of poor access to fertilizers and modern farming technologies[11,16]. In such contexts, new roads—or road improvements such as paving—could increase access to agricultural supplies and markets, facilitating production increases and lowering post-harvest crop losses[13,17]. As such accessible areas tend to sustain more prosperous rural livelihoods, they may also act as 'magnets', attracting colonists away from environmentally vulnerable frontier areas, such as the margins of forests[17,18]. In this way, improving transportation in suitable areas could help to concentrate and improve agricultural production, raising farm yields[11,13] while potentially promoting land sparing for nature conservation[19].

Despite the pivotal role that roads have in human land-use, efforts to plan and zone roads are extremely inadequate. First, although roads increasingly dominate much of Earth's land surface (Fig. 1), many roads are unmapped, especially in developing nations; in the Brazilian Amazon, for example, the total length of unofficial or illegal roads is nearly triple that of official roads[20]. Second, environmental-impact assessments often place the burden of proof on road opponents[21,22], who rarely have sufficient information on rare species, biological resources and ecosystem services[23] needed to determine the actual environmental costs of roads. Third, many road assessments are limited in scope[4,22], focusing only on the direct effects of road building while ignoring its critical indirect effects, such as promoting deforestation, fires, poaching and land speculation. Finally, because there is no strategic, proactive system for zoning roads globally, road projects must be assessed with little information on their broader context (see the 2013 report on high-risk road development by the Conservation Strategy Fund; http://conservation-strategy.org/sites/default/files/field-file/CSFPolicyBrief_14_english_1.pdf). This increases the burden on road planners and evaluators, who are being swamped by the unprecedented pace of contemporary road expansion[2–7,11,15,20].

For these reasons, we devised a 'global roadmap' to identify areas in which roads or road improvements are likely to have major costs or benefits. The map has two components: an environmental-values layer that estimates the natural importance of ecosystems, and a road-benefits layer that estimates the potential for increased agricultural production, in part via new or improved roads. Combining these two layers allows us to identify areas where roads or road upgrades could have large potential benefits, areas where road building should be avoided wherever possible, and conflict areas where their potential costs and benefits are both sizeable.

We created the environmental-values layer (Fig. 2a) by integrating global data sets on three classes of parameters: biodiversity (number of threatened terrestrial-vertebrate species, estimated number of plant species per ecoregion); key wilderness habitats (G200 terrestrial ecoregions, important bird areas and endemic bird areas, biodiversity hotspots, frontier forests, high-biodiversity wilderness areas); and carbon storage and climate-regulation services of the local ecosystem (see Methods and Supplementary Figs 1–11). Values for each class were equally weighted, rescaled (range: 0–1) and then averaged to produce the environmental-values layer. Regions that scored highly on this layer include wet and humid tropical and subtropical forests, Mediterranean ecosystems, wildlife-rich savanna woodlands in South America and Africa, many islands, certain mountain ranges, and some higher-latitude forests, among others.

The road-benefits layer (Fig. 2b) identifies areas where new roads or road improvements could potentially help to improve agricultural production. Like the environmental-values layer, it is a relative index (range: 0–1). In general terms, areas that score highly on this layer have been largely converted to agriculture (and thus have little native vegetation remaining), are relatively low-yielding despite having soils and climates

[1]Centre for Tropical Environmental and Sustainability Science, and College of Marine and Environmental Sciences, James Cook University, Cairns, Queensland 4878, Australia. [2]Kenyir Research Institute, Universiti Malaya Terengganu, 21030 Kuala Terengganu, Malaysia. [3]Institute on the Environment, and Department of Ecology, Evolution, and Behavior, University of Minnesota, Saint Paul, Minnesota 55108, USA. [4]Center for the Environment, Harvard University, Cambridge, Massachusetts 02138, USA. [5]Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK. [6]Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK. [7]Australian Research Centre for Urban Ecology, and School of Botany, University of Melbourne, Melbourne, Victoria 3010, Australia. [8]Conservation Strategy Fund, 663-2300 Curridabat, San José, Costa Rica.

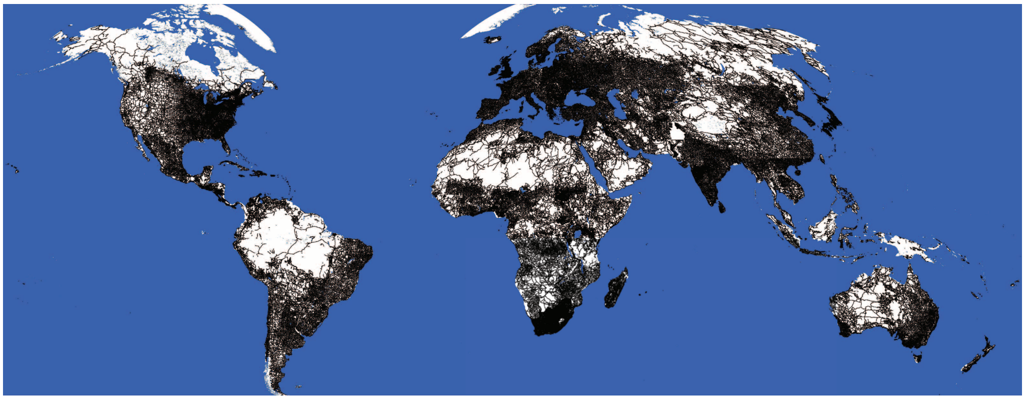


Figure 1 | The distribution of major roads globally. Roads are indicated in black; white areas lack mapped roads. The quality of road maps varies greatly among nations, with many smaller and unofficial roads remaining unmapped. We generated this map using data from the integrated gROADS database (http://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1; accessed 7 June 2014); Center for International Earth Science Information Network - CIESIN - Columbia University, and Information Technology Outreach Services - ITOS - University of Georgia. 2013. Global Roads Open Access Data Set, Version 1 (gROADSv1). Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4VD6WCT.

broadly suitable for agriculture, are not so distant from urban markets that crop-transportation costs would be prohibitive even with new or improved roads, and are expected to see large future increases in agricultural production to meet projected food or export demands (see Methods and Supplementary Figs 12–16 for details of how these data sets were integrated). All continents have regions that score highly, including parts of south Asia, east and southeast Asia, West and East Africa, central Eurasia, west-central North America, Central America and Mexico, and the Atlantic region of South America.

We classified each of the environmental-values (Fig. 2a) and road-benefits (Fig. 2b) layers into deciles and then cross-tabulated them to generate 100 unique colour combinations (see Supplementary Information for details). In this scheme, green-shaded areas are where road building would have relatively high environmental costs and only modest potential benefits for agriculture. Red-shaded areas are the opposite, with high potential to increase agricultural production and lower scores on the environmental-values axis. Black and dark-shaded areas are 'conflict zones' with high values on both axes, whereas white and light-shaded areas are lower priorities for both environment and agriculture.

On top of this scheme we overlaid polygons for 177,857 protected areas (Supplementary Fig. 17) globally, using available data from the World Database on Protected Areas (http://www.wdpa.org). Protected areas

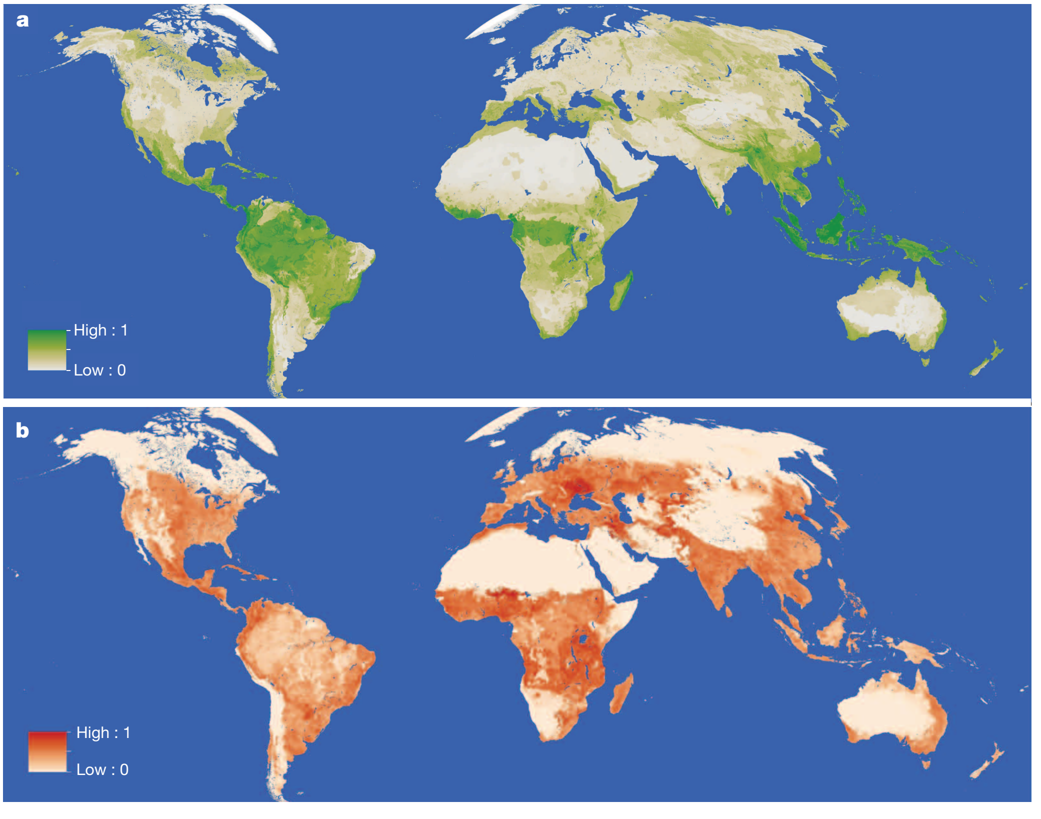


Figure 2 | The environmental-values and road-benefits layers. a, b, The environmental-values layer (**a**) integrates data on terrestrial biodiversity, key habitats, wilderness, and environmental services. The road-benefits layer (**b**) shows areas broadly suitable for agricultural intensification, where new roads or road improvements could potentially promote increased production. See Supplementary Information for data sources.

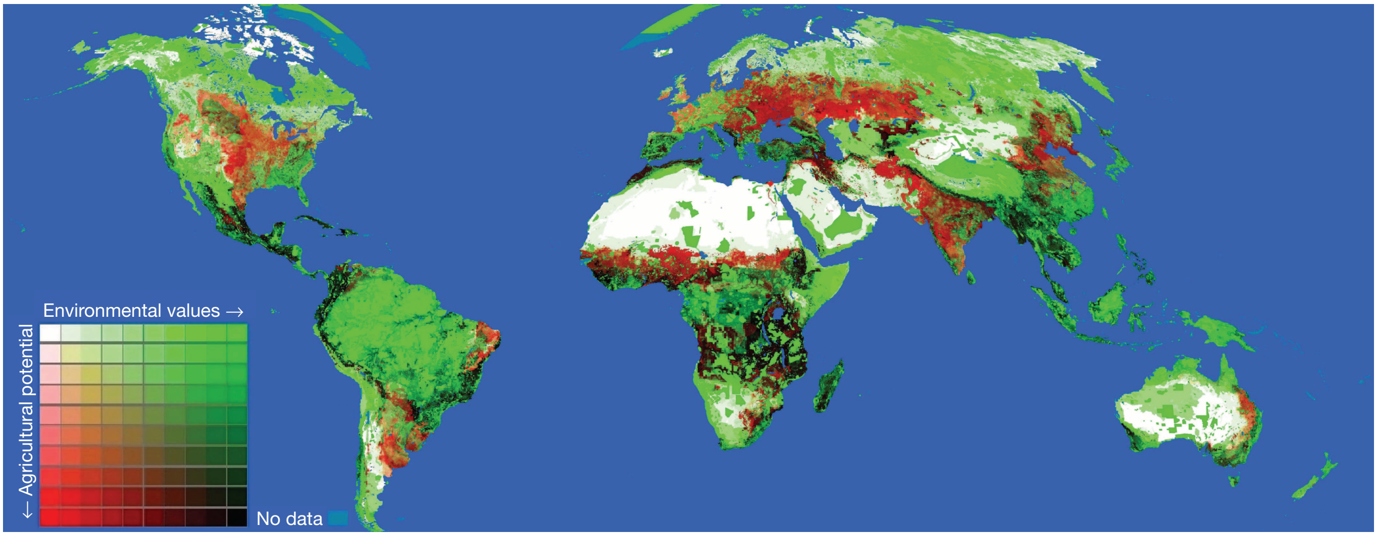


Figure 3 | A global roadmap. Shown are priority road-free areas (green shades), priority agricultural areas (red shades), conflict areas (dark shades), and lower-priority areas (light shades). Values of the environmental-values and road-benefits layers are each divided into deciles, yielding 100 unique colour combinations. See Supplementary Information for details and data sources.

were zoned fully green because we judged that they should be free of new roads wherever possible, given that roads can facilitate illegal activities such as poaching, encroachment, and vehicle-related road-kill of wildlife[2–4] that are contrary to the goals of protected-area management[24,25].

The resulting global roadmap (Fig. 3) attempts to portray key relative risks and rewards of road building for each 1-km^2 pixel on Earth's land surface. In broad terms, our map illustrates the enormous potential for environmental loss and degradation as a result of contemporary road expansion (Table 1 and Supplementary Fig. 18). Roads are currently proliferating or planned in many areas categorized as having high environmental values but only modest agricultural potential, such as the Amazon Basin, parts of the Asia-Pacific region, and higher-latitude forests in the Northern Hemisphere.

The roadmap also reveals extensive conflict areas (Fig. 3), where environmental and agricultural values are both high, particularly in Sub-Saharan Africa, Madagascar, Central America, the Mediterranean, southeast and south-central Asia, the Andes, and the Atlantic region of South America. Conflict zones often occur in regions with rapid population growth, high species endemism, or both. In total, 1.97 billion hectares (16.5% of global land area) fall into conflict areas (Table 1). Land-use pressures in such regions are mounting rapidly; it has been estimated that, unless current agricultural yields markedly improve, approximately 1 billion hectares of additional farming and grazing land will be needed by 2050 to meet projected food demands[9], with extensive additional lands converted for production of biofuels[26].

However, our road-planning scheme also suggests that many areas could be targeted for agricultural production increases with relatively modest environmental costs. Such areas include expanses of the Indian subcontinent, central Eurasia, the Irano-Anatolian region, and African Sahel, among others (Fig. 3). In total, 1.46 billion hectares of land (12.3% of global land area) is zoned red (Table 1), suggesting that there is considerable potential on every continent to increase agricultural production, by raising yields on existing farming and grazing land.

Although improved roads or other transportation can facilitate agricultural yield increases[11,13,17,18], additional measures—such as investments in improved farming methods, fertilizers and, where appropriate, irrigation—will also be essential. A particular challenge will be devising strategies to help developing nations with exceptional environmental values, such as Madagascar and Indonesia (Fig. 2a), to meet pressing economic and food-production needs while limiting the environmental costs of rapid road development. For such nations, international payments for ecosystem services, ecotourism, and sustainable harvesting of native production forests could potentially help to balance economic and environmental priorities[27]. A further priority when planning road and agricultural investments is to consider how factors such as inter-annual weather variability or projected future climate change could impact on crop yields[28].

The global roadmap we created underscores the potential benefits and need for strategic road planning, but actual road planning will be undertaken at smaller national or regional scales. For this, we created more detailed maps that show finer-scale features (for example, Extended Data Fig. 1). These maps and their components are freely available (http://global-roadmap.org) and can be combined with additional data, such as more detailed information on topography, soils, existing croplands and local road networks, to facilitate road planning.

Integrating local information is important because the drivers and environmental impacts of road construction will vary in different contexts. For example, in arable, largely road-free areas of East Africa (Fig. 4a), new roads driven by a burgeoning mining boom[11,29] could provoke major land-use changes and habitat loss. Yet expanding roads from timber and mining operations could also have large impacts in Siberia (Fig. 4b), even

Table 1 | Percentages of seven geographical regions that fall into four broad categories on the global roadmap

Zone	Africa	Asia	Australia	Europe	North and Central America	South America	Oceania	Global
Conserve	29.03	45.69	34.21	26.44	47.39	66.28	95.29	46.31
Agriculture	7.93	12.44	3.63	32.92	11.35	6.83	0.23	12.29
Conflict	24.75	14.87	7.01	9.10	8.70	15.74	0.58	16.54
Low-tension	38.30	27.00	55.15	31.54	32.55	11.14	3.89	32.67
Total area	29,805	44,174	7,693	9,670	23,395	17,662	412	132,811

Data on the total areas of each region are given in $km^2 \times 10^3$. 'Conserve' zones are where road building would have relatively high environmental costs (above-median environmental values; Fig. 2a) and modest potential agricultural benefits (below-median road-benefits values; Fig. 2b). 'Agriculture' zones have the opposite attributes (above-median road-benefits values and below-median environmental values). 'Conflict' zones have both above-median environmental values and above-median road-benefits values, whereas 'low-tension' zones are lower priorities for both environment and agriculture (with below-median environmental and road-benefits values). See Supplementary Fig. 18 for a map of these zones.

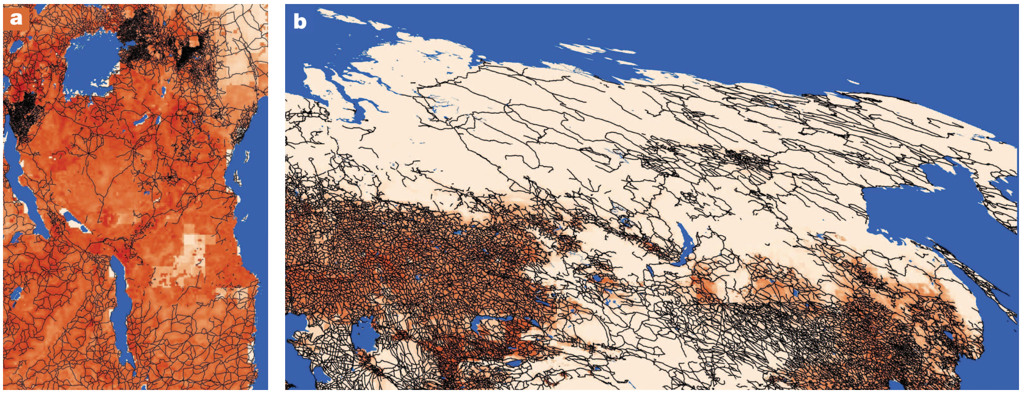


Figure 4 | Mapped roads overlaid onto the roads-benefits layer. **a**, **b**, In eastern Africa (**a**) and Siberia (**b**), roads are rapidly expanding into relatively road-free areas, but for different reasons. Narrow black lines indicate mapped roads. In both regions, areas with darker-red colours have greater agricultural potential than those with lighter colours. See Supplementary Information for data sources.

though agricultural potential is limited, by promoting forest fires and clearing[14]. In general, we expect road impacts to be lowest in unproductive, arid regions, moderate in carbon-rich ecosystems such as higher-latitude forests, and most damaging in species- and carbon-rich ecosystems such as tropical forests, particularly where few roads currently exist.

We see our global road-mapping scheme as a working model—an important first step towards strategic road planning to reduce environmental damage—that can be downscaled and tailored for particular circumstances. We believe such proactive planning should be a central element of any discussion about road expansion and associated land-use zoning[13,30]. Given that the total length of new roads anticipated by mid-century[1] would encircle the Earth more than 600 times, there is little time to lose.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 May; accepted 28 July 2014.
Published online 27 August 2014.

1. Dulac, J. *Global Land Transport Infrastructure Requirements: Estimating Road and Railway Infrastructure Capacity and Costs to 2050* (International Energy Agency, 2013).
2. Laurance, W. F. *et al.* The future of the Brazilian Amazon. *Science* **291,** 438–439 (2001).
3. Blake, S. *et al.* Forest elephant crisis in the Congo Basin. *PLoS Biol.* **5,** e111 (2007).
4. Laurance, W. F., Goosem, M. & Laurance, S. G. Impacts of roads and linear clearings on tropical forests. *Trends Ecol. Evol.* **24,** 659–669 (2009).
5. Adeney, J. M., Christensen, N. & Pimm, S. L. Reserves protect against deforestation fires in the Amazon. *PLoS ONE* **4,** e5014 (2009).
6. Fearnside, P. M. & Graça, P. BR-319: Brazil's Manaus-Porto Velho Highway and the potential impact of linking the arc of deforestation to central Amazonia. *Environ. Manage.* **38,** 705–716 (2006).
7. Forman, R. T. T. *et al. Road Ecology: Science and Solutions* (Island Press, 2003).
8. Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R. & Polasky, S. Agricultural sustainability and intensive production practices. *Nature* **418,** 671–677 (2002).
9. Tilman, D. *et al.* Forecasting agriculturally driven global environmental change. *Science* **292,** 281–284 (2001).
10. Perz, S. G. *et al.* Regional integration and local change: road paving, community connectivity and social-ecological resilience in a tri-national frontier, southwestern Amazonia. *Reg. Environ. Change* **12,** 35–53 (2012).
11. Weng, L. *et al.* Mineral industries, growth corridors and agricultural development in Africa. *Glob. Food Security* **2,** 195–202 (2013).
12. Boakes, E. H., Mace, G. M., McGowan, P. J. K. & Fuller, R. A. Extreme contagion in global habitat clearance. *Proc. R. Soc. Lond. B* **277,** 1081–1085 (2010).
13. Laurance, W. F. & Balmford, A. A global map for road building. *Nature* **495,** 308–309 (2013).
14. Bradshaw, C. J. A., Warkentin, I. G. & Sodhi, N. S. Urgent preservation of boreal carbon stocks and biodiversity. *Trends Ecol. Evol.* **24,** 541–548 (2009).
15. Laporte, N. T., Stabach, J. A., Grosch, R., Lin, T. S. & Goetz, S. J. Expansion of industrial logging in central Africa. *Science* **316,** 1451 (2007).
16. Mueller, N. D. *et al.* Closing yield gaps through nutrient and water management. *Nature* **490,** 254–257 (2012).
17. Weinhold, D. & Reis, E. Transportation costs and the spatial distribution of land use in the Brazilian Amazon. *Glob. Environ. Change* **18,** 54–68 (2008).
18. Rudel, T. K., DeFries, R., Asner, G. P. & Laurance, W. F. Changing drivers of deforestation and new opportunities for conservation. *Conserv. Biol.* **23,** 1396–1405 (2009).
19. Phalan, B., Onial, M., Balmford, A. & Green, R. E. Reconciling food production and biodiversity conservation: Land sharing and land sparing compared. *Science* **333,** 1289–1291 (2011).
20. Barber, C. P., Cochrane, M. A., Souza, C. M. Jr & Laurance, W. F. Roads, deforestation, and the mitigating effect of protected areas in the Amazon. *Biol. Conserv.* **177,** 203–209 (2014).
21. Gullett, W. Environmental impact assessment and the precautionary principle: Legislating caution in environmental protection. *Australas. J. Environ. Manage.* **5,** 146–158 (1998).
22. Laurance, W. F. Forest destruction: the road to ruin. *New Sci.* **194,** 25 (2007) http://www.newscientist.com/article/mg19426075.600-forest-destruction-the-road-to-ruin.html.
23. Lawrence, D. P. *Environmental Impact Assessment: Practical Solutions to Recurrent Problems* (John Wiley & Sons, 2003).
24. Laurance, W. F. *et al.* Averting biodiversity collapse in tropical forest protected areas. *Nature* **489,** 290–294 (2012).
25. Caro, T., Dobson, A., Marshall, A. J. & Peres, C. A. Compromise solutions between conservation and road building in the tropics. *Curr. Biol.* **24,** R722–R725 (2014).
26. Warner, E. *et al.* Modeling biofuel expansion effects on land use change dynamics. *Environ. Res. Lett.* **8,** 015003 (2013).
27. Campbell, W. B. & López Ortíz, S. (eds) *Integrating Agriculture, Ecotourism, and Conservation: Examples from the Field* (Springer, 2011).
28. Challinor, A. J. *et al.* A meta-analysis of crop yield under climate change and adaptation. *Nature Clim. Chang.* **4,** 287–291 (2014).
29. Edwards, D. P. *et al.* Mining and the African environment. *Conserv. Lett.* **7,** 302–311 (2014).
30. Balmford, A., Green, R. & Phalan, B. What conservationists need to know about farming. *Proc. R. Soc. Lond. B* **279,** 2714–2724 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank T. Brooks, S. Butchart, J. Geldmann, S. Goosem, C. Mendenhall, N. Pares, S. Pimm, U. Srinivasan, N. Velho, and two anonymous referees for comments and feedback. The Australian Research Council provided support.

Author Contributions W.F.L. and A.B. initially conceived the study, and W.F.L. coordinated its design, analysis, and manuscript preparation. G.R.C. and S.S. conducted the spatial analyses; C.S.O., N.D.M., O.V., G.R.C., S.S. and B.P. generated or collated key datasets; and M.G., D.P.E., R.V.D.R. and I.B.A. provided ideas and critical feedback.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to W.F.L. (bill.laurance@jcu.edu.au).

doi:10.1038/nature13451

The evolution of the placenta drives a shift in sexual selection in livebearing fish

B. J. A. Pollux[1,2], R. W. Meredith[1,3], M. S. Springer[1], T. Garland[1] & D. N. Reznick[1]

The evolution of the placenta from a non-placental ancestor causes a shift of maternal investment from pre- to post-fertilization, creating a venue for parent–offspring conflicts during pregnancy[1–4]. Theory predicts that the rise of these conflicts should drive a shift from a reliance on pre-copulatory female mate choice to polyandry in conjunction with post-zygotic mechanisms of sexual selection[2]. This hypothesis has not yet been empirically tested. Here we apply comparative methods to test a key prediction of this hypothesis, which is that the evolution of placentation is associated with reduced pre-copulatory female mate choice. We exploit a unique quality of the livebearing fish family Poeciliidae: placentas have repeatedly evolved or been lost, creating diversity among closely related lineages in the presence or absence of placentation[5,6]. We show that post-zygotic maternal provisioning by means of a placenta is associated with the absence of bright coloration, courtship behaviour and exaggerated ornamental display traits in males. Furthermore, we found that males of placental species have smaller bodies and longer genitalia, which facilitate sneak or coercive mating and, hence, circumvents female choice. Moreover, we demonstrate that post-zygotic maternal provisioning correlates with superfetation, a female reproductive adaptation that may result in polyandry through the formation of temporally overlapping, mixed-paternity litters. Our results suggest that the emergence of prenatal conflict during the evolution of the placenta correlates with a suite of phenotypic and behavioural male traits that is associated with a reduced reliance on pre-copulatory female mate choice.

Viviparity creates a venue for parent–offspring conflicts *in utero*[7] caused by a fundamental discord between mothers and developing embryos over the level of maternal investment during pregnancy. Females are selected to maximize their lifetime reproductive success by optimizing the allocation to each offspring while individual offspring are selected to demand a greater investment from the mother than is optimal for her to provide[1–4]. The ensuing evolutionary dynamics of perpetual adaptation and counter-adaptation between mother and developing embryo are hypothesized to be the driving force behind a rapid divergence in the genomic, developmental and physiological details of the placenta[1,3,6].

A central tenet of the parent–offspring conflict theory is that offspring must be able to manipulate the transfer of resources[1,7]. However, not all viviparous taxa have this capability[4–6]. Lecithotrophic viviparous species lack placentas and allocate all resources necessary for embryo development to the eggs before fertilization. This limits the potential for viviparity-driven conflict, because maternal investment pre-dates the expression of the paternal genome[1,2,4,6]. The evolution of the placenta from a non-placental lecithotrophic ancestor causes a shift in maternal investment from pre- to post-fertilization[5,6], offering embryos the opportunity to influence maternal investment throughout gestation[6,8,9]. This creates the potential for genomic conflicts, the magnitude of which depend on the extent of post-zygotic investment[1–4,6].

Theory predicts that the emergence of genomic conflicts, early in the evolution of the placenta, should drive a shift from a reliance on pre-copulatory mate choice to increasing levels of polyandry in conjunction with post-zygotic mechanisms of sexual selection[2]. Lecithotrophic species produce large, 'costly' (that is, fully provisioned) eggs[5,6], gaining most reproductive benefits by carefully selecting suitable mates based on phenotype or behaviour[2]. These females, however, run the risk of mating with genetically inferior (for example, closely related or dishonestly signalling) males, because genetically incompatible males are generally not discernable at the phenotypic level[10]. Placental females may reduce these risks by producing tiny, inexpensive eggs and creating large mixed-paternity litters by mating with multiple males. They may then rely on the expression of the paternal genomes to induce differential patterns of post-zygotic maternal investment among the embryos and, in extreme cases, divert resources from genetically defective (incompatible) to viable embryos[1–4,6,11].

Here we apply comparative methods to examine potential conflict-driven shifts in sexual selection associated with the evolution of post-fertilization maternal provisioning within the livebearing fish family Poeciliidae (order Cyprinodontiformes). This family presents a unique opportunity, because (1) it contains closely related species that differ markedly in the degree and timing of maternal provisioning, ranging from strict pre-zygotic yolk provisioning to extreme levels of post-zygotic investment associated with integrated maternal and fetal tissues specialized for nutrient transfer (that is, placentas[5,6]); (2) placentas were lost or evolved multiple times independently[5,6]; and (3) there is great interspecific variation in reproductive traits associated with pre-copulatory sexual selection, including caudal swords, enlarged dorsal fins and bright coloration in males[12–14]. Furthermore, a number of lineages have evolved the ability to carry multiple, temporally overlapping litters that are fertilized at different points in time (that is, superfetation[6,13,14]). In mammals, superfetation facilitates the formation of mixed-paternity litters (polyandry)[15–17]. Finally, molecular and experimental studies suggest that prenatal genomic conflicts occur in this family[18] and can result in differential patterns of post-zygotic maternal investment between developing embryos[19].

If substantial post-fertilization maternal provisioning intensifies fetal–maternal conflict[1,3] causing reduced female reliance on pre-copulatory cues in mate choice[2], then males of species with extensive post-fertilization maternal provisioning should display less developed, or the absence of, traits that facilitate female mate choice before copulation. Such traits include sexual dichromatism, courtship behaviour or ornaments. Moreover, if superfetation facilitates multiple paternity[15–17], then species with relatively high levels of post-fertilization provisioning should also have a higher probability of having superfetation[20,21]. Finally, copulation is known to incur costs to the females (for example, physical injury, reduced feeding opportunity, increased risk of predation and/or sexually transmitted diseases)[13,14]. If substantial post-fertilization maternal provisioning coincides with an increase in the frequency of polyandry, the ensuing sexual conflict should drive the evolution of female (resistance) traits that reduce the costs associated with superfluous mating attempts and, at the same time, male traits that enhance male mating success in the face of female resistance[14]. In Poeciliidae, a smaller male size relative to female size and an increase in gonopodium length (the male copulatory organ) increases the reproductive success of males

[1]Department of Biology, University of California, Riverside, California 92521, USA. [2]Experimental Zoology Group, Wageningen University, 6708 WD Wageningen, the Netherlands. [3]Department of Biology and Molecular Biology, Montclair State University, Montclair, New Jersey 07043, USA.

during sneaky or coercive copulation, which enables males to circumvent female choice[14,22–24]. We thus predict that males of species with a relatively higher post-zygotic maternal investment should display relatively smaller body sizes and longer gonopodia.

To test these hypotheses, we first quantified the degree of post-fertilization maternal provisioning for each species with the 'matrotrophy index', which is the estimated dry mass of the offspring at birth divided by the dry mass of the egg at fertilization. The matrotrophy index provides an objective, dimensionless measure of the degree of post-fertilization maternal provisioning that presents a proxy for the level of placentation[5,6]. Lecithotrophic species have matrotrophy index values of less than 1, because embryos lose dry mass during gestation[5,6]. Placentotrophic species have matrotrophy index values greater than 1, because post-fertilization maternal provisioning causes growth during development[5,6]. We employed a well-resolved phylogeny to test for predicted evolutionary shifts in sexual selection with Bayesian tests

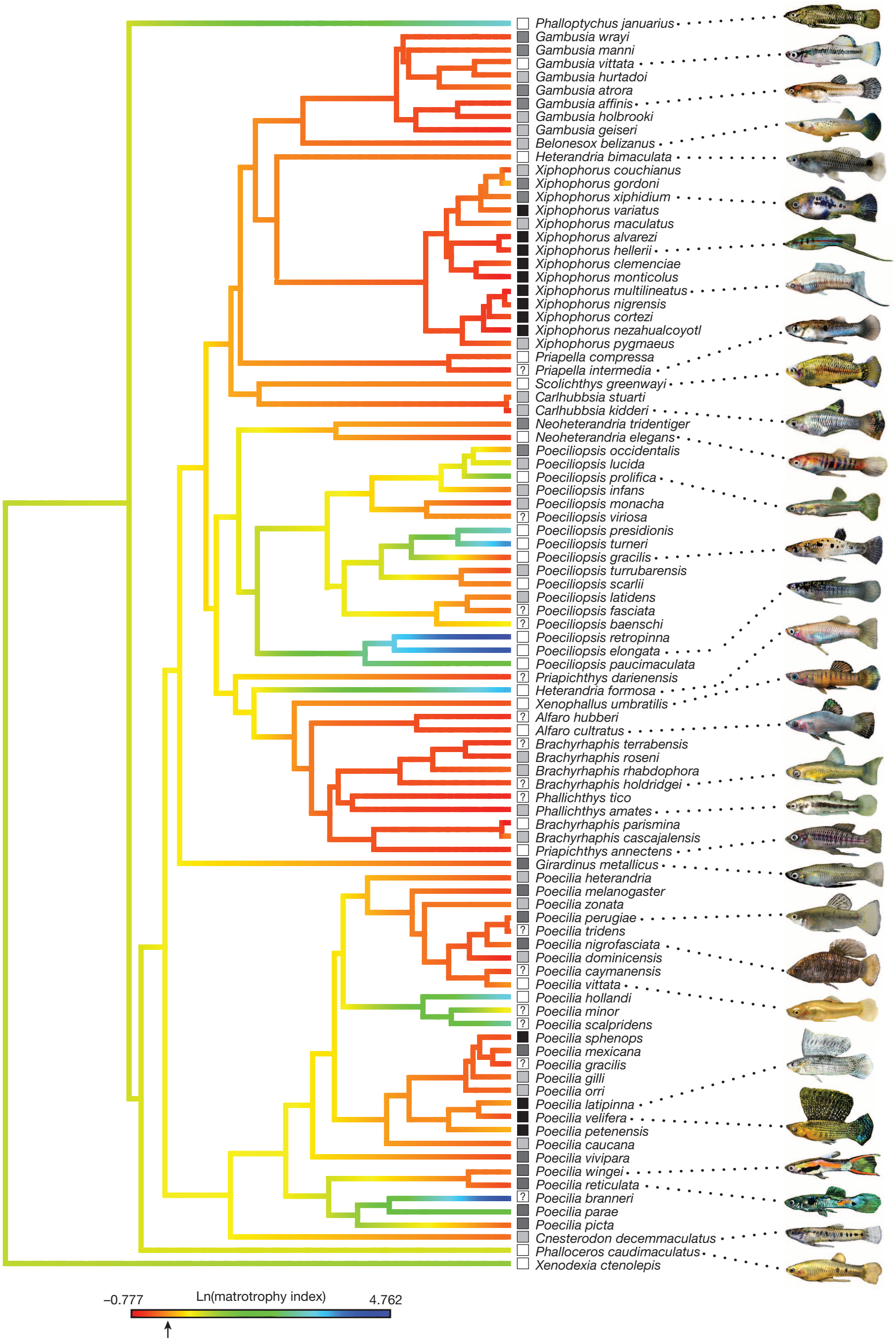


Figure 1 | Phylogenetic tree showing relationships among 94 species of the fish family Poeciliidae. Boxes at the terminal ends of the branches are coded according to the male sexual selection index: black = 3, dark grey = 2, light grey = 1 and white = 0; the boxed question mark indicates incomplete information (Supplementary Table 1). Branch colours depict a maximum likelihood reconstruction of maternal provisioning for log-transformed matrotrophy indices. The ancestral reconstruction was performed with phytools[30] and a Brownian motion model of trait evolution. The arrow on the scale bar corresponds to a matrotrophy index value of 1.0, which indicates the division between lecithotrophic and placentotrophic species. In agreement with previous analyses[5,6], the ancestral reconstruction suggests a complex history for the evolution of placentotrophy. The current analysis suggests that the common ancestor of the family has a placenta and that there were multiple losses and gains of placentation within the family. A caveat is that the single egg-layer within Poeciliidae (*Tomeurus gracilis*) was excluded from the analysis. The inclusion of this taxon, along with outgroups that contain both livebearers and egg-layers, may yield different results for the evolutionary history of placentotrophy within Poeciliidae.

of correlated trait evolution and phylogenetic linear and logistic regressions that explain the variation in sexually selected traits as a function of matrotrophy index (Fig. 1).

Bayesian tests show that matrotrophy index, and sexually dimorphic coloration (dichromatism) and courtship behaviour, respectively, have evolved in a correlated fashion (log(Bayes factor) = 13.308 and 3.438, respectively). Phylogenetic logistic regressions further show that both traits are negatively correlated to matrotrophy index ($b_1 = -0.614$, $P = 0.007$ and $b_1 = -0.763$, $P = 0.007$, respectively), indicating that these are significantly less likely to be found in placental lineages (Fig. 2a, b). The presence of exaggerated male display traits (that is, enlarged dorsal fins and filamentous extensions on upper maxillae in the genus *Poecilia* or extension of ventral part of the caudal fin to form a sword in *Xiphophorus*) is also negatively correlated with matrotrophy index, but this trend is not significant after correcting for phylogeny ($b_1 = -0.413$, $P = 0.429$; log(Bayes factor) = 0.663; Fig. 2c). The sexual selection index, defined as the summed presence of these three male traits (dichromatism, courtship behaviour and ornamental display traits), decreases significantly with increasing matrotrophy index (phylogenetic generalized least-squares regression: $F_{77} = 5.836$, $P = 0.018$; log(Bayes factor) = 2.320), indicating that lecithotrophic males have significantly more traits to facilitate female choice before copulation than highly placental species (Fig. 2d).

Superfetation is strongly correlated with matrotrophy index (phylogenetic logistic regression: $b_1 = 0.776$, $P < 0.001$; log(Bayes factor) = 25.730; Fig. 2e), indicating that placental species are more likely to have it. The relative length of the gonopodium is positively correlated with matrotrophy index (phylogenetic generalized least-squares regression: $F_{89} = 6.379$, $P = 0.013$; log(Bayes factor) = 4.214; Fig. 2f), demonstrating a strong association between longer genitalia and the presence of post-fertilization maternal provisioning. The size dimorphism index is also positively correlated with matrotrophy index, both for body weight ($F_{76} = 18.869$, $P < 0.001$; log(Bayes factor) = 6.664; Fig. 2g) and standard length ($F_{87} = 29.753$, $P < 0.001$; log(Bayes factor) = 4.948; Fig. 2h), indicating that the difference in body size between males and females is larger in lineages with higher levels of post-zygotic maternal investment. This increase is caused by a decrease in male size ($\log_{10}$(male wet mass): $F_{81} = 3.493$, $P = 0.065$, log(Bayes factor) = 3.676; $\log_{10}$(male standard length): $F_{89} = 2.022$, $P = 0.158$, log(Bayes factor) = 1.310; Fig. 2i, j blue lines) in association with increasing matrotrophy index, rather than an increase in female size ($\log_{10}$(female wet mass): $F_{76} = 0.021$, $P = 0.886$. log(Bayes factor) = 0.122; $\log_{10}$(female standard length): $F_{87} = 0.002$, $P = 0.962$, log(Bayes factor) = 0.298; Fig. 2i, j red lines).

Our findings yield three important insights. First, male traits that facilitate pre-copulatory female mate choice are less well developed in placental lineages. This is supported by patterns within individual clades. In the northern clade of the genus *Poeciliopsis*[5], males from lecithotrophic species are melanic whereas males from derived placental species have the same coloration as females. In the subgenus *Micropoecilia* of *Poecilia*, males belonging to the lecithotrophic clade are far more intensely coloured than the males in the derived placental clade, suggesting that here too sexually dimorphic coloration is disappearing. Extreme male ornamental display traits used during courtship are only found in lecithotrophic clades (*Xiphophorus* and subgenus *Mollienesia* of *Poecilia*) and are notably absent in placental species. Since sexual selection in Poeciliidae is influenced by pre-copulatory cues[12–14], these results suggest that phenotype- or behaviourally based female mate choice is of greater importance in lecithotrophic species than in species with substantial post-zygotic maternal provisioning.

Second, male traits that help circumvent female mate choice during sneak or coercive mating are more developed in placental species. These findings concur with the theory that sexual conflict can result in the evolution of sexual dimorphism[25,26] and rapid phenotypic divergence in genitalia[27,28]. In poeciliids, large males and short genitals are associated with courtship behaviour aimed at attracting cooperative females[12–14]. Smaller males and longer genitalia are associated with sneak copulation, the small size allowing males to approach females from behind without being detected and enabling them to manoeuvre more easily when inserting the gonopodium into the gonoduct of uncooperative females, while longer gonopodia enable a more efficient sperm transfer during unsolicited matings[12–14,22–24].

Third, the degree of post-zygotic maternal provisioning is strongly correlated with superfetation, which is found in all placental lineages save for *Phalloceros caudimaculatus* and the subgenus *Pamphorichthys* of *Poecilia*. This reproductive adaptation is thought to diminish the

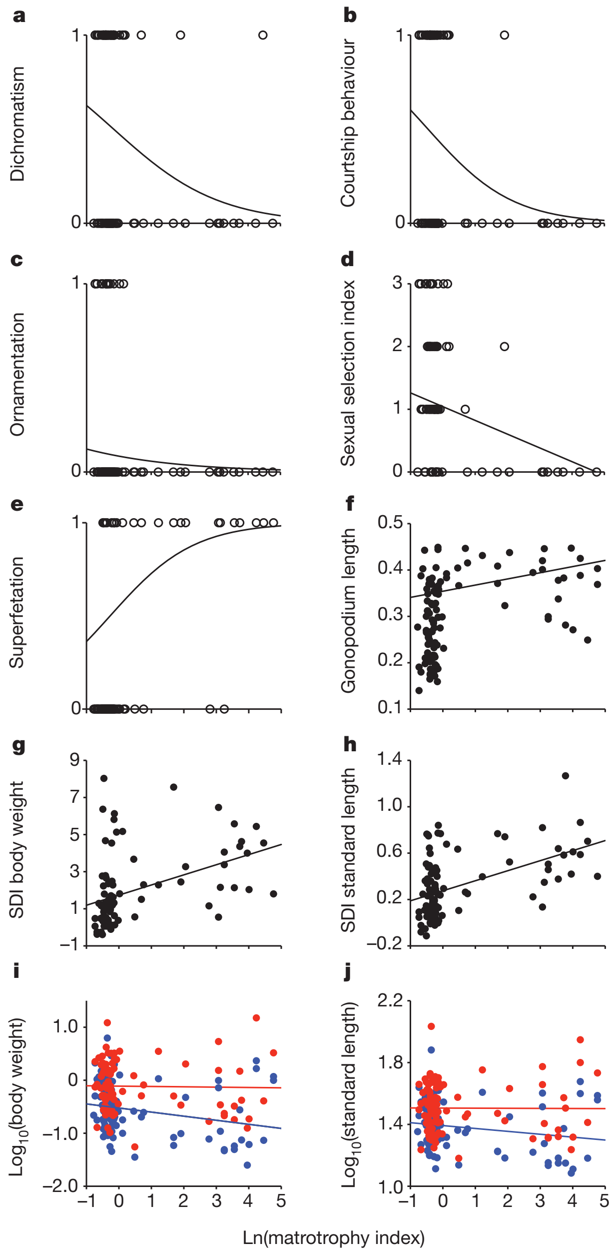


Figure 2 | Phylogenetic logistic and linear regressions. The regressions evaluate the effect of the natural-log-transformed matrotrophy index on (**a**) dichromatism ($n = 94$ taxa), (**b**) courtship behaviour ($n = 79$), (**c**) ornamental male display traits ($n = 94$), (**d**) sexual selection index (defined as the total number of male traits present ranging from 0 (none of the three traits present) to maximum 3 (all three traits present); $n = 79$), (**e**) superfetation ($n = 92$), (**f**) relative gonopodium length ($n = 107$ taxa), (**g**) size dimorphism index (SDI) for body weight ($n = 87$), (**h**) size dimorphism index for standard length ($n = 100$), (**i**) $\log_{10}$-transformed male (blue dots and line, $n = 99$) and female (red dots and line, $n = 87$) body weight, and (**j**) $\log_{10}$-transformed male (blue dots and line, $n = 107$) and female (red dots and line, $n = 100$) standard length.

probability of a single male monopolizing an entire litter by fertilizing all embryos. Instead, by dividing offspring into multiple, smaller temporally overlapping litters, each fertilized at different points in time, and by using sperm derived from the most recent mating event ('last male sperm precedence'[13,14]), superfetation increases a female's likelihood of creating multiple-paternity litters[15–17].

Prior research has shown that male coloration, courtship behaviour and ornamental display traits play an important role in pre-copulatory female mate choice[13–14], that small male size and long genitalia facilitate sneak copulation (a strategy that circumvents female choice)[13,14,22–24] and that superfetation facilitates the formation of mixed-paternity litters (polyandry)[15–17]. The correlation of these traits with the level of post-zygotic maternal provisioning provides support for an association between the placenta and weaker pre-copulatory mate choice. What remains to be shown is that this relationship is causal and is associated with an increase in multiple paternities. Our study provides the first empirical evidence concurring with the hypothesis[2] that the rise of parent–offspring conflicts during the evolutionary transition from pre- to post-zygotic maternal provisioning correlates with a shift in sexual selection. This study will help to understand the elusive consequences of viviparity-driven conflict and may advance our knowledge about the evolution of reproductive traits in other viviparous lineages that evolved placentas, since all share the same potential for genomic conflict.

METHODS SUMMARY

The maximum likelihood phylogeny was constructed using RAxML 7.0.4 (ref. 29). Different phylogenetic comparative approaches were used to test for correlated trait evolution.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 August 2013; accepted 2 May 2014.

Published online 9 July; corrected online 10 September 2014 (see full-text HTML version for details).

1. Haig, D. Genetic conflicts in human pregnancy. *Q. Rev. Biol.* **68,** 495–532 (1993).
2. Zeh, D. W. & Zeh, J. A. Reproductive mode and speciation: the viviparity-driven conflict hypothesis. *Bioessays* **22,** 938–946 (2000).
3. Wilkins, J. R. & Haig, D. What good is genomic imprinting: the function of parent-specific gene expression. *Nature Rev. Genet.* **4,** 359–368 (2003).
4. Crespi, B. & Semeniuk, C. Parent-offspring conflict in the evolution of vertebrate reproductive mode. *Am. Nat.* **163,** 635–653 (2004).
5. Reznick, D. N., Mateos, M. & Springer, M. S. Independent origins and rapid evolution of the placenta in the fish genus *Poeciliopsis. Science* **298,** 1018–1020 (2002).
6. Pollux, B. J. A., Pires, M. N., Banet, A. I. & Reznick, D. N. The evolution of placentas in the fish family Poeciliidae – an empirical study of macroevolution. *Annu. Rev. Ecol. Evol. Syst.* **40,** 271–289 (2009).
7. Trivers, R. L. Parent-offspring conflict. *Am. Zool.* **14,** 249–264 (1974).
8. Banet, A. I., Au, A. G. & Reznick, D. N. Is mom in charge? Implications of resource provisioning on the evolution of the placenta. *Evolution* **64,** 3172–3182 (2010).
9. Pollux, B. J. A. & Reznick, D. N. Matrotrophy limits a female's ability to adaptively adjust offspring size and fecundity in fluctuating environments. *Funct. Ecol.* **25,** 747–756 (2011).
10. Zeh, J. A. & Zeh, D. W. Toward a new sexual selection paradigm: polyandry, conflict and incompatibility. *Ethology* **109,** 929–950 (2003).
11. Haig, D. Brood reduction and optimal parental investment when offspring differ in quality. *Am. Nat.* **136,** 550–556 (1990).
12. Bisazza, A. Male competition, female mate choice and sexual size dimorphism in poeciliid fishes. *Mar. Behav. Physiol.* **23,** 257–286 (1993).
13. Meffe, G. K. & Snelson, F. F. Jr. (eds). *Ecology and Evolution of Livebearing Fishes (Poeciliidae)* (Prentice Hall, 1989).
14. Evans, J. P., Pilastro, A. & Schlupp, I. (eds) *Ecology and Evolution of Poeciliid Fishes* (Univ. Chicago Press, 2011).
15. Shackelford, R. M. Superfetation in the ranch mink. *Am. Nat.* **86,** 311–319 (1952).
16. Yamaguchi, N., Dugdale, H. L. & MacDonald, D. W. Female receptivity, embryonic diapause, and superfetation in the European badger (*Meles meles*): implications for the reproductive tactics of males and females. *Q. Rev. Biol.* **81,** 34–48 (2006).
17. Dugdale, H. L., MacDonald, D. W., Pope, L. C. & Burke, T. Polygynandry, extra-group paternity and multiple-paternity litters in European badger (*Meles meles*) social groups. *Mol. Ecol.* **16,** 5294–5306 (2007).
18. O'Neill, M. J. *et al.* Ancient and continuing Darwinian selection on insulin-like growth factor II in placental fishes. *Proc. Natl Acad. Sci. USA* **104,** 12404–12409 (2007).
19. Schrader, M. & Travis, J. Testing the viviparity-driven-conflict hypothesis: parent-offspring conflict and the evolution of reproductive isolation in a poeciliid fish. *Am. Nat.* **172,** 806–817 (2008).
20. Coleman, S. W., Harlin-Cognato, A. & Jones, A. G. Reproductive isolation, reproductive mode, and sexual selection: Empirical tests of the viviparity-driven conflict hypothesis. *Am. Nat.* **173,** 291–303 (2009).
21. Schrader, M. & Travis, J. Variation in offspring size with birth order in placental fish: A role for asymmetric sibling competition? *Evolution* **66,** 272–279 (2012).
22. Bisazza, A. & Marin, G. Sexual selection and sexual size dimorphism in the eastern mosquitofish *Gambusia holbrooki* (Pisces Poeciliidae). *Ethol. Ecol. Evol.* **7,** 169–183 (1995).
23. Pilastro, A., Giacomello, E. & Bisazza, A. Sexual selection for small size in male mosquitofish (*Gambusia holbrooki*). *Proc. R. Soc. Lond. B* **264,** 1125–1129 (1997).
24. Evans, J. P. *et al.* Intraspecific evidence from guppies for correlated patterns of male and female genital trait diversification. *Proc. R. Soc. B* **278,** 2611–2620 (2011).
25. Arnqvist, G. & Rowe, L. Antagonistic coevolution between the sexes in a group of insects. *Nature* **415,** 787–789 (2002).
26. Bonduriansky, R. & Chenoweth, S. F. Intralocus sexual conflict. *Trends Ecol. Evol.* **24,** 280–288 (2009).
27. Arnqvist, G. Comparative evidence for the evolution of genitalia by sexual selection. *Nature* **393,** 784–786 (1998).
28. Hosken, D. J. & Stockley, P. Sexual selection and genital evolution. *Trends Ecol. Evol.* **19,** 88–93 (2004).
29. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22,** 2688–2690 (2006).
30. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3,** 217–223 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank all individuals and institutions who provided samples for this study (Rehoboth Aquatics, H. Bart Jr, R. Davis, D. Fromm, J. de Greef, H. Hieronimus, B. Hobbs, T. Hrbek, J. Johnson, B. Kohler, J. Langenhammer, C. Li, J. Lundberg, M. Mateos, A. Meyer, D. Nelson, L. Page, L. Parenti, M. Sabaj Pérez, R. Robins, R. de Ruiter, S. Schaefer, M. Schartl, J. Sparks, M. Stiassny, J. Travis, J. Trexler and J. Williams); L. Rowe and A. Furness for discussions and reading the manuscript; and C. Oufiero and M. Banet for help in collecting part of the data. This study was supported by Rubicon grant 825.07.017 of the Netherlands Organisation for Scientific Research and Marie Curie – IIF grant 299048 of the European Union to B.J.A.P. and grant DEB0416085 of the US National Science Foundation to D.N.R. and M.S.S.

Author Contributions B.J.A.P. and D.N.R. designed the study, D.N.R. quantified the matrotrophy indices, R.W.M. and M.S.S. constructed the molecular phylogeny, T.G. taught B.J.A.P. how to do phylogenetic regression and aided in the preliminary analysis of the data, and B.J.A.P. measured the morphological traits, performed the final analyses of the data and wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.J.A.P. (b.pollux@gmail.com; bart.pollux@wur.nl) or D.N.R. (david.reznick@ucr.edu).

doi:10.1038/nature13449

Innate immune sensing of bacterial modifications of Rho GTPases by the Pyrin inflammasome

Hao Xu[1]*, Jieling Yang[1,2]*, Wenqing Gao[1], Lin Li[1], Peng Li[1], Li Zhang[1], Yi-Nan Gong[1], Xiaolan Peng[1], Jianzhong Jeff Xi[3], She Chen[1], Fengchao Wang[1] & Feng Shao[1,2,4]

Cytosolic inflammasome complexes mediated by a pattern recognition receptor (PRR) defend against pathogen infection by activating caspase 1. Pyrin, a candidate PRR, can bind to the inflammasome adaptor ASC to form a caspase 1-activating complex[1,2]. Mutations in the Pyrin-encoding gene, *MEFV*, cause a human autoinflammatory disease known as familial Mediterranean fever[3–5]. Despite important roles in immunity and disease, the physiological function of Pyrin remains unknown. Here we show that Pyrin mediates caspase 1 inflammasome activation in response to Rho-glucosylation activity of cytotoxin TcdB[6–8], a major virulence factor of *Clostridium difficile*, which causes most cases of nosocomial diarrhoea. The glucosyltransferase-inactive TcdB mutant loses the inflammasome-stimulating activity. Other Rho-inactivating toxins, including FIC-domain adenylyltransferases (*Vibrio parahaemolyticus* VopS and *Histophilus somni* IbpA) and *Clostridium botulinum* ADP-ribosylating C3 toxin, can also biochemically activate the Pyrin inflammasome in their enzymatic activity-dependent manner. These toxins all target the Rho subfamily and modify a switch-I residue. We further demonstrate that *Burkholderia cenocepacia* inactivates RHOA by deamidating Asn 41, also in the switch-I region, and thereby triggers Pyrin inflammasome activation, both of which require the bacterial type VI secretion system (T6SS). Loss of the Pyrin inflammasome causes elevated intra-macrophage growth of *B. cenocepacia* and diminished lung inflammation in mice. Thus, Pyrin functions to sense pathogen modification and inactivation of Rho GTPases, representing a new paradigm in mammalian innate immunity.

C. difficile is the major cause of hospital-acquired infectious diarrhoea and antibiotic-associated pseudomembranous colitis. The major virulence factors of *C. difficile* are two secreted protein toxins (TcdA and TcdB)[7,8]. TcdA/B and the related *Clostridium sordellii* lethal toxin (TcsL) belong to the large clostridial glucosylating cytotoxin family that inactivates members of Rho and/or Ras-family small GTPases by monoglucosylating a threonine residue critical for GTP binding[6]. Recent studies indicate that TcdA/B can activate the inflammasome[9,10]. Consistently, recombinant TcdB triggered robust caspase 1 activation, interleukin (IL)-1β production and pyroptosis in primary bone marrow-derived macrophages (BMDMs) (Fig. 1a and Extended Data Fig. 1a–d). TcsL, sharing 85% sequence homology to TcdB, showed no such activities. As expected, both TcdB and TcsL caused cell rounding, which did not occur with the glucosyltransferase-deficient TcdB(W102A/D288N) and TcsL(D286N/D288N) mutants[6] (referred to as TcdB^{m} and TcsL^{m}, respectively) (Extended Data Fig. 2a). TcdB^{m} could not induce caspase 1 inflammasome activation (Fig. 1a and Extended Data Fig. 1a–d). Thus, the Rho-glucosylating activity of TcdB but not TcsL triggers inflammasome activation in mouse macrophages.

The PYRIN-CARD domain adaptor ASC is critical for caspase 1 activation mediated by a PYRIN-domain PRR. Consistent with previous studies[9,10], inflammasome activation in *Asc* (also known as *Pycard*)$^{-/-}$ BMDMs was resistant to TcdB stimulation (Fig. 1a and Extended Data Fig. 1c–d). By contrast, *Nlrp3*$^{-/-}$ *Nlrc4*$^{-/-}$ and *Aim2*$^{-/-}$ BMDMs remained sensitive to TcdB. As a control, enterohaemorrhagic *Escherichia coli* (EHEC) type III secretion needle protein EprI induced little IL-1β production in *Nlrp3*$^{-/-}$ *Nlrc4*$^{-/-}$ and *Asc*$^{-/-}$ BMDMs, whereas pyroptosis was only diminished in *Nlrp3*$^{-/-}$ *Nlrc4*$^{-/-}$ BMDMs. Deletion of *Nod1/2*, involved in sensing bacterial activation of Rho GTPases in the NF-κB pathway[11,12], did not affect TcdB-induced caspase 1 activation (Extended Data Fig. 1e–g). TcdB induced similar caspase 1 activation in BMDMs from different mouse inbred strains (Extended Data

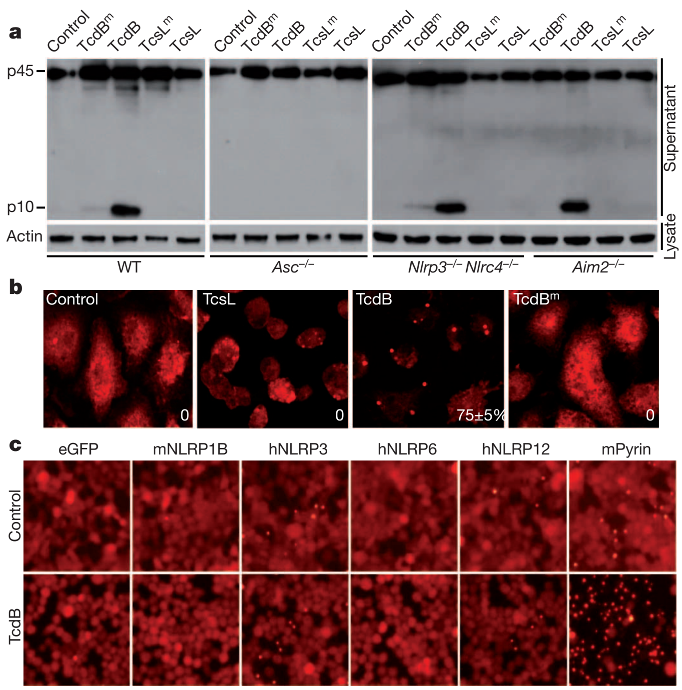


Figure 1 | Inflammasome activation by TcdB and identification of Pyrin as the candidate immune sensor. **a**, Assays of inflammasome activation by TcdB and TcsL in BMDMs from wild-type (WT, C57BL/6) or indicated knockout mice. Macrophage supernatants were collected for anti-caspase 1 immunoblotting (p45, pro-caspase 1; p10, mature caspase 1). **b**, Anti-ASC immunostaining of TcdB- and TcsL-stimulated BMDMs. Percentages of cells showing the ASC foci are marked. **c**, Assays of different PYRIN- or CARD-domain proteins in supporting TcdB-induced ASC foci formation in 293T cells stably expressing RFP–ASC. mNLRP1B, mouse NLRP1B; hNLRP3/6/12, human NLRP3/6/12; mPyrin, mouse Pyrin. TcdB^{m} and TcsL^{m} denote the glucosyltransferase-deficient TcdB(W102A/D288N) and TcsL(D286N/D288N) mutants, respectively. Data in all panels are representative of at least three repetitions.

[1]National Institute of Biological Sciences, Beijing 102206, China. [2]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China. [3]Department of Biomedical Engineering, College of Engineering, Peking University, Beijing 100871, China. [4]National Institute of Biological Sciences, Beijing, Collaborative Innovation Center for Cancer Medicine, Beijing 102206, China.
*These authors contributed equally to this work.

Fig. 1h). TcdB, but not TcdB^m or TcsL, rendered endogenous ASC aggregation (foci formation) in BMDMs (Fig. 1b). Thus, TcdB activates an ASC-containing inflammasome independently of NLRP3, NLRC4, AIM2 and NOD1/2.

Red fluorescent protein (RFP)–ASC foci formation reconstituted in 293T cells was used to screen for the upstream PRR protein in sensing TcdB. Of a total of ten PYRIN-domain proteins tested (NLRP2, NLRP3, NLRP5, NLRP6, NLRP7, NLRP8, NLRP9, NLRP11, NLRP12 and Pyrin), only Pyrin supported TcdB-induced RFP–ASC foci formation (Fig. 1c and data not shown). Mutation of Pyrin-encoding *MEFV* causes the human autoinflammatory disorder familial Mediterranean fever (FMF)[3,4]. Consistently, Pyrin has been shown to interact with ASC through their amino-terminal PYRIN domains to promote caspase 1 activation *in vitro*[1,2,13]. FMF-causing mutations are gain-of-function and the disease-like symptom in the knock-in mice results from ASC-dependent excessive IL-1β release[5]. Pyrin is primarily expressed in monocytes and dendritic cells. We noticed that, in contrast to primary BMDMs, immortalized mouse BMDMs (iBMDMs), primary bone marrow-derived dendritic cells and immortalized mouse DC2.4 dendritic cells showed no inflammasome response to TcdB stimulation (Fig. 2a). *Mefv* messenger RNA was not detectable in these three cell types (Fig. 2b). Stable expression of mouse or human Pyrin but not NLRP3 in DC2.4 cells rendered robust inflammasome responses to TcdB without increasing the sensitivity to the NAIP inflammasome agonist (*Shigella flexneri* type III secretion needle protein MxiH) and the NLRP3 inflammasome agonist lipopolysaccharide (LPS) plus nigericin (Fig. 2c and Extended Data Fig. 3a–c). Human THP-1 monocytes showed inflammasome activation upon TcdB stimulation whereas U937 cells did not (Extended Data Fig. 3d–f). Consistently, Pyrin expression in U937 cells was six times lower than that in THP-1 cells (Extended Data Fig. 3g). Ectopic expression of Pyrin in U937 cells regained TcdB-induced caspase 1 activation (Extended Data Fig. 3d, f). In DC2.4 cells, TcdB but not TcdB^m converted enhanced green fluorescent protein (eGFP)–Pyrin from a dispersed localization into large cytosolic aggregates (Fig. 2d). eGFP–Pyrin indeed co-aggregated with endogenous ASC, which did not occur with poly(dA:dT) that triggered ASC aggregation through AIM2 inflammasome activation (Fig. 2d and Extended Data Fig. 3h).

Small interference RNA (siRNA) knockdown of Pyrin in primary BMDMs inhibited TcdB-induced caspase 1 activation (Fig. 3a). The knockdown efficiency of three different siRNAs correlated with their inhibition on caspase 1 activation (Fig. 3a, b). We generated Pyrin-deficient mice using transcription activator-like effector nuclease (TALEN)-based genome editing technology and analysed five independent homozygous mutant mice (F_1) (Extended Data Fig. 4a, b). When BMDMs from *Mefv*$^{-/-}$ strains (F_1-1 and F_1-2) were stimulated with TcdB, *Salmonella typhimurium* flagellin (FliC), or LPS plus nigericin, only TcdB-induced inflammasome activation was diminished (Fig. 3c–e). The results were confirmed in BMDMs from two additional *Mefv*$^{-/-}$ strains (F_1-3 and F_1-4) (Extended Data Fig. 4c). Thus, Pyrin is required for TcdB-induced inflammasome activation.

Both TcdB and TcsL caused cell rounding owing to an inactivating modification of Rho GTPases (Extended Data Figs 2a and 5a–c), suggesting that actin cytoskeleton disruption is unlikely to be the cause of Pyrin activation. Supporting this idea, cytochalasin D and the actin cross-linking domain (ACD) from *Vibrio cholerae* RTX toxin[14] induced severe cell rounding but no evident caspase 1 activation (Extended Data Fig. 6a, b). The RID domain of *Vibrio* RTX toxin that induces Rho-GTP hydrolysis without covalent modification[15] did not stimulate inflammasome activation in primary BMDMs and Pyrin-reconstituted DC2.4 and 293T cells (Extended Data Fig. 6c–e). Thus, Pyrin specifically responds to toxin-catalysed inactivating modifications of Rho GTPases.

Rho GTPases are frequent targets of bacterial pathogens[16]. Recent studies identify several FIC-domain Rho-adenylylating effectors: *V. parahaemolyticus* VopS adenylylates the same threonine as TcdB[17] (Thr 37 in RHOA; Thr 35 in Rac/Cdc42) and the two FIC domains in *H. somni* IbpA modify a nearby tyrosine (Tyr 34 in RHOA; Tyr 32 in Rac/Cdc42)[18,19]. Similar to TcdB, the three FIC domains target the Rho subfamily and Rac/Cdc42 but not other Ras-superfamily members[18]. When purified VopS or IbpA-Fic1/2 was delivered into primary BMDMs using an anthrax lethal factor N-terminal domain (LFn)-fusion strategy, apparent cell rounding and covalent modification of RHOA occurred (Extended Data Figs 2b and 5d), both of which were abolished by mutation of the catalytic histidine (H348A in VopS, H3295A in IbpA-Fic1 and H3717A in IbpA-Fic2). All three FIC-domain proteins, but not their catalytic histidine mutants, induced evident caspase 1 inflammasome activation, which required *Asc* but not *Nlrp3* and *Nlrc4* (Extended Data Fig. 7a, b). In addition, similar to TcdB, VopS and IbpA-Fic1/2-induced inflammasome activation was diminished in *Mefv*$^{-/-}$ BMDMs (Fig. 4a and Extended Data Fig. 7c, d). VopS-positive *V. parahaemolyticus* induced robust caspase 1 activation in Pyrin-reconstituted DC2.4 cells but a weak Pyrin-independent response in primary BMDMs (Extended Data Fig. 7e, f). Biochemical activation of the Pyrin inflammasome observed with recombinant VopS and IbpA-Fic1/2 further supports that Pyrin senses pathogen modification of Rho GTPases.

TcdB and FIC-domain effectors modify both the Rho subfamily (RHOA/B/C) and Rac/Cdc42, whereas TcsL only targets Rac/Cdc42, Ras and Ras-related Ral/Rap GTPases[20,21]. We confirmed that RHOA was inactivated in TcdB but not TcsL-stimulated BMDMs whereas cellular Rac and Cdc42 were inactivated by both toxins (Extended Data Fig. 5b, c). *C. botulinum* C3 toxin, the first and most established Rho-modifying toxin[22], is highly selective for RHOA/B/C[16]. C3 ADP-ribosylates Asn 41 in RHOA, causing a slower migration on an SDS–PAGE gel (Extended Data Fig. 8a). C3 and TcdB modifications are mutually exclusive owing to the physical proximity between Asn 41 and Thr 37 (in RHOA). Taking advantage of this property, we confirmed that endogenous

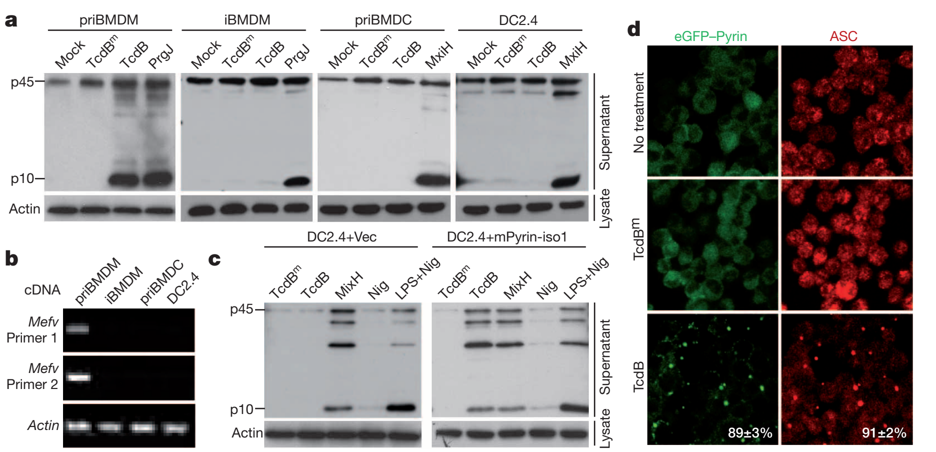


Figure 2 | Pyrin mediates TcdB-induced inflammasome activation. **a**, **b**, Profiling of the sensitivity of different macrophage/dendritic cells to TcdB and their Pyrin expression. Mouse primary BMDMs (priBMDM), immortalized BMDMs (iBMDM), primary bone marrow-derived dendritic cells (priBMDC) and DC2.4 cells were stimulated as indicated. PrgJ and MxiH were delivered by the LFn-PA (protective antigen) system. *Mefv* mRNA levels in **b** were measured by reverse transcriptase (RT)–PCR. **c**, **d**, TcdB-induced inflammasome activation in Pyrin-complemented DC2.4 cells. DC2.4 cells harbouring a vector (Vec) or mouse Pyrin isoform 1 (mPyrin-iso1) or stably expressing eGFP–Pyrin were stimulated as indicated (Nig, nigericin). Caspase 1 immunoblots are shown in **a** and **c**; eGFP and anti-ASC fluorescence images are in **d**. Data in all panels are representative from at least three repetitions.

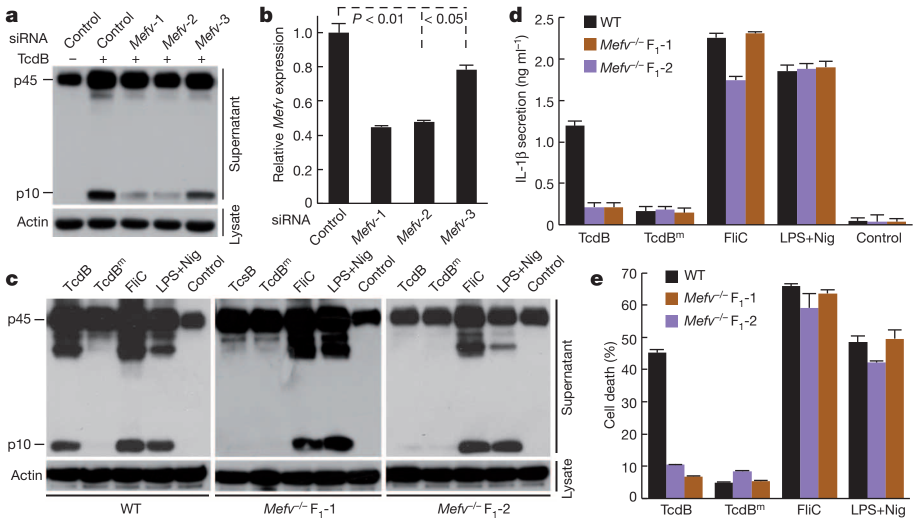


Figure 3 | Pyrin is required for TcdB-induced inflammasome activation. **a**, **b**, Effects of *Mefv* knockdown on TcdB-induced caspase 1 activation in primary BMDMs. The knockdown efficiency of three independent *Mefv*-targeting siRNAs (*Mefv*-1/2/3) was measured by quantitative (q)RT–PCR in **b** ($n = 3$, mean ± s.d.; P value, Student's t-test). **c**–**e**, Effects of *Mefv* knockout on TcdB-induced inflammasome activation. $Mefv^{-/-}$ mice were generated by TALEN-mediated targeting (Extended Data Fig. 4). FliC (*S. typhimurium* flagellin) was delivered by the LFn-PA system. F_1-1 and F_1-2 are two independent $Mefv^{-/-}$ F_1 lines. Representative caspase 1 immunoblots from at least three repetitions are in **a** and **c**; ELISA assays of IL-1β release and cell death measured by LDH release are in **d** and **e**, respectively ($n = 3$; mean ± s.d.).

RHOA in BMDMs was modified by TcdB but not TcsL (Extended Data Fig. 5a) whereas C3 toxin modified RHOA but not Cdc42 (Extended Data Fig. 8b). Consistently, RHOA but not Rac/Cdc42 was inactivated in C3-intoxicated macrophages, contrary to TcsL stimulation (Extended Data Fig. 5b, c). Similar to TcdB, C3 triggered extensive caspase 1 activation and IL-β secretion in primary BMDMs (Extended Data Fig. 8c, d). The catalytically inactive $C3^{m}$ mutant (Q212A/E214A) that did not modify cellular RHOA (Extended Data Fig. 8a) induced no inflammasome activation (Extended Data Fig. 8c, d). C3-induced inflammasome activation was independent of NLRP3 and NLRC4, but abolished in $Asc^{-/-}$ and $Mefv^{-/-}$ BMDMs (Fig. 4b and Extended Data Fig. 8c–e). A TcsL H17-C variant (replacing α17 in TcsL with the corresponding helix from TcdB) could recognize RHOA and catalyse the same modification as TcdB[23] (Fig. 4c). TcsL H17-C also induced caspase 1 activation, which was diminished in $Mefv^{-/-}$ BMDMs (Fig. 4d). The modification sites by TcdB, FIC-domain effectors and C3 toxin (Thr 37, Tyr 34 and Asn 41 in RHOA) are all in the GTPase switch-I region. These together strongly suggest that switch-I modification of RHOA or other Rho-subfamily members but not Rac/Cdc42 induces Pyrin inflammasome activation.

In a 293T cell reconstitution system (stably expressing Pyrin and RFP–ASC), when expression of RHOA/B/C was individually knocked down by specific siRNAs (Extended Data Fig. 9a), percentages of cells developing RFP–ASC foci upon TcdB stimulation were not significantly decreased (Extended Data Fig. 9b). However, triple knockdown of RHOA/B/C markedly reduced TcdB-induced RFP–ASC foci formation (Fig. 4e and Extended Data Fig. 9b). Similar results were obtained with another independent set of siRNAs (Extended Data Fig. 9c, d). The deficient ASC foci formation in the triple-knockdown cells could be restored by stable expression of RNAi-resistant RHOA whereas the modification-site mutant of RHOA(T37A) showed no rescue effects (Fig. 4f). These results establish the requirement of Rho modification for toxin-induced Pyrin activation and also indicate a functional redundancy of RHOA/B/C. In both 293T and DC2.4 cells, no Pyrin–Rho interaction could be detected even in the presence of TcdB (Extended

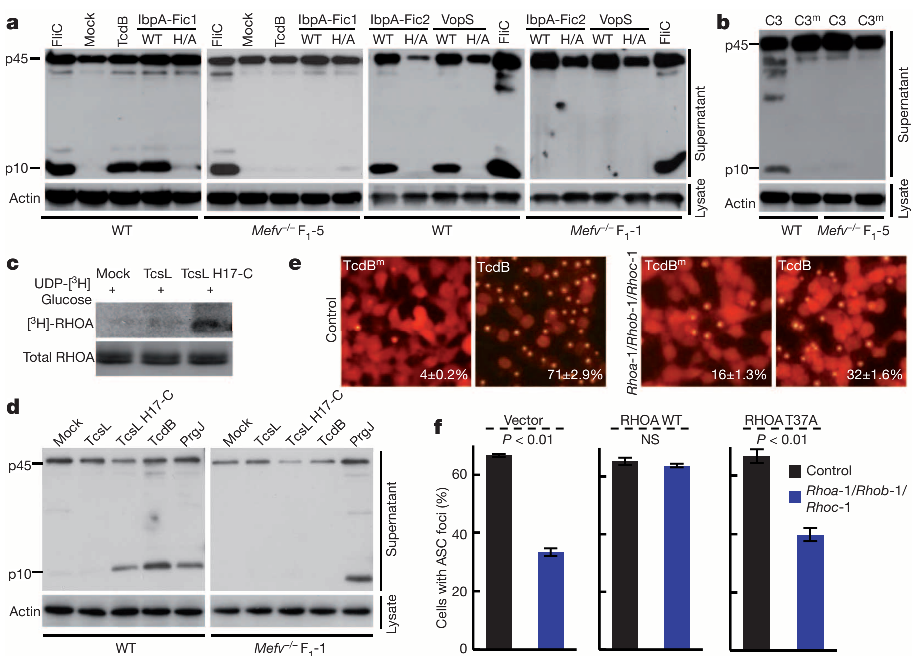


Figure 4 | Switch-I modification of the Rho subfamily accounts for Pyrin activation by TcdB and other Rho-modifying toxins. **a**, **b**, Caspase 1 activation by *V. parahaemolyticus* VopS, *H. somni* IbpA (IbpA-Fic1/2) and *C. botulinum* C3 toxin (delivered into BMDMs by the LFn-PA system) and effects of *Mefv* knockout. H/A, mutation of the FIC-domain catalytic histidine; $C3^{m}$, the catalytically inactive Q212A/E214A mutant. **c**, *In vitro* glucosylation of RHOA by TcsL and TcsL H17-C. Shown are ^{3}H autoradiograph and immunoblot of RHOA. **d**, Wild-type or $Mefv^{-/-}$ BMDMs were treated with TcsL H17-C or indicated control stimuli. **e**, **f**, Requirement of the Rho subfamily and its modification for TcdB-induced inflammasome activation. 293T cells stably expressing Pyrin and RFP–ASC were transfected with control or a pool of three siRNAs targeting RHOA, B and C, respectively (*Rhoa*-1, *Rhob*-1 and *Rhoc*-1). The cells were additionally stable-transfected with a vector (left), or RNAi-resistant RHOA wild type (WT; middle) or T37A (right) before Rho knockdown and TcdB stimulation in **f**. Percentages of cells showing the ASC foci (marked on the fluorescence images in **e**) are mean ± s.d. ($n = 3$) (P value, Student's t-test; NS, not significant). Representative caspase 1 immunoblots from at least three repetitions are shown in **a**, **b** and **d**.

Data Fig. 9e, f), suggesting a different mechanism of Pyrin sensing of Rho modification/inactivation from the NOD2–RIP2 axis that directly binds to Rho-GTP[12,24,25].

Intracellular *B. cenocepacia* causes fatal chronic lung infection in immunocompromised individuals. *B. cenocepacia* inactivates Rho and disrupts actin cytoskeleton in a T6SS-dependent manner[26,27], which was confirmed in DC2.4 cells (Extended Data Fig. 5e). RHOA in wild-type *B. cenocepacia*-infected but not T6SS-deficent Δ*hcp* strain-infected macrophages was resistant to C3 modification, suggesting a modification by *B. cenocepacia* (Fig. 5a). Mass spectrometric analysis of Flag–RHOA purified from uninfected or *B. cenocepacia*-infected DC2.4 cells was therefore performed. Among all identified tryptic peptides covering 97% of the RHOA sequence, modification of one peptide, $_{28}$DQFPEVYVPTVFENYVADIEVDGK$_{51}$, was found in *B. cenocepacia*-infected cells; tandem mass spectrometry revealed that Asn 41 within the peptide was deamidated into an aspartate (Fig. 5b). The extracted ion chromatogram showed that more than 90% of Asn 41-containing peptides recovered from wild-type infection were deamidated whereas Asn 41-deamidated peptide was barely detectable in uninfected or Δ*hcp*-infected cells (Fig. 5c). Thus, *B. cenocepacia* infection induces RHOA deamidation on Asn 41 in a T6SS-dependent manner.

Asn 41 is the same site modified by C3 toxin, indicating a role of Pyrin in defending *B. cenocepacia* infection. A recent study reports that Pyrin knockdown decreases IL-1β secretion in *B. cenocepacia*-infected human monocytes[28]. We observed efficient caspase 1 processing in wild-type *B. cenocepacia* but not Δ*hcp*-infected mouse BMDMs (Extended Data Fig. 10a). Similarly to that observed with TcdB, *Nlrp3*$^{-/-}$ *Nlrc4*$^{-/-}$ and *Aim2*$^{-/-}$ BMDMs showed intact inflammasome response to *B. cenocepacia* (Extended Data Fig. 10b, c), whereas little caspase 1 activation and IL-1β secretion were detected in infected *Mefv*$^{-/-}$ and *Asc*$^{-/-}$ BMDMs (Fig. 5d and Extended Data Fig. 10b–e). *Mefv*$^{-/-}$ did not affect *S. typhimurium*-induced inflammasome activation. Furthering these genetic analyses, re-expression of Pyrin in DC2.4 and U937 cells restored caspase 1 inflammasome activation by *B. cenocepacia* (but not EHEC, which harbours no Rho-modifying effectors), which remained dependent upon the T6SS (Extended Data Fig. 10f–i). In the 293T cell

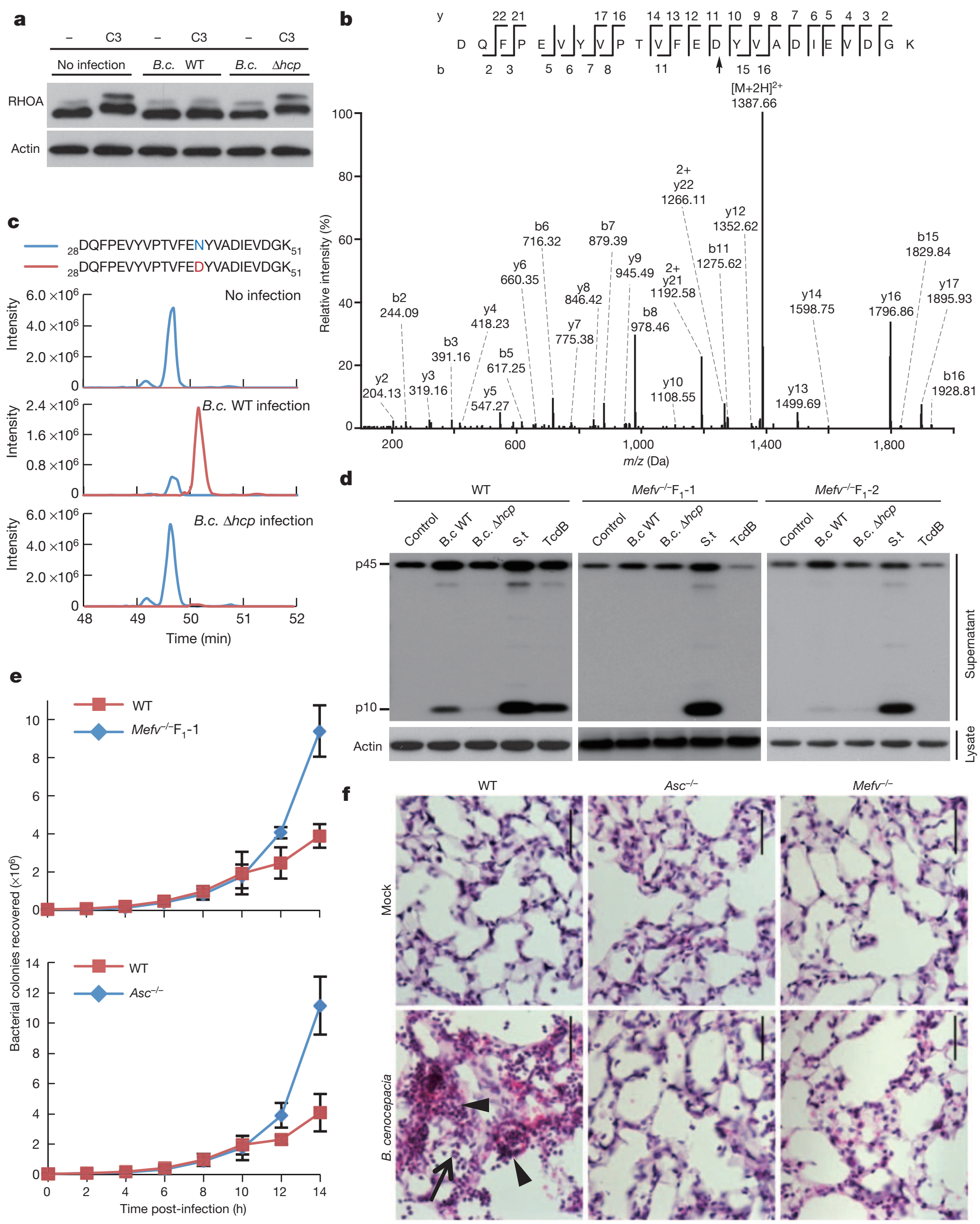


Figure 5 | T6SS-dependent Asn 41 deamidation of RHOA by *B. cenocepacia* activates the Pyrin inflammasome. **a**, *B. cenocepacia* infection-induced RHOA modification in BMDMs probed by the sensitivity to *in vitro* C3 modification. **b**, Tandem mass spectrum of a tryptic peptide of Flag–RHOA from wild-type *B. cenocepacia*-infected DC2.4 cells. The b- and y-type product ions are marked and illustrated along the peptide sequence shown on top of the spectrum. Arrowhead, the deamidated Asn 41. **c**, Flag–RHOA purified from uninfected or *B. cenocepacia*-infected DC2.4 cells was analysed by mass spectrometry. Shown are extracted ion chromatograms of the Asn 41-containing peptide. **d**, Effects of *Mefv* knockout on *B. cenocepacia*-induced caspase 1 activation in primary BMDMs. **e**, Effects of *Mefv* and *Asc* knockout on *B. cenocepacia* intracellular replication. The numbers of bacterial colonies recovered from BMDM lysates ($n = 3$; mean ± s.d.) are plotted against the time duration of infection. **f**, Indicated mice were infected with *B. cenocepacia* or a buffer control (mock). Representative haematoxylin & eosin staining of the lung sections is shown. Black triangle and arrow mark inflammatory cell infiltration and intra-alveolar leukocytes, respectively. Scale bars, 50 μm. *B. c.* WT and Δ*hcp* indicate wild-type and T6SS-deficient *B. cenocepacia* J2315, respectively. Data in all panels are representative of at least three repetitions.

reconstitution system, overexpression of deamidated (N41D) but not wild-type RHOA could drive RFP–ASC foci formation when endogenous RHOA/B/C were simultaneously knocked down (Extended Data Fig. 10j). This confirms that *B. cenocepacia*-induced RHOA deamidation accounts for Pyrin inflammasome activation. Furthermore, *B. cenocepacia* replication in *Mefv*$^{-/-}$ and *Asc*$^{-/-}$ BMDMs was comparable but much higher than that in wild-type BMDMs (Fig. 5e). *B. cenocepacia*-infected mice developed strong lung inflammation, marked by inflammatory cell infiltration, appearance of intra-alveolar leukocytes and disruption of the normal lung architecture. These responses were severely attenuated in the lungs of infected *Mefv*$^{-/-}$ and *Asc*$^{-/-}$ mice (Fig. 5f). Thus, the Pyrin inflammasome is a critical for immune defence against *B. cenocepacia* by sensing bacterial T6SS-induced deamidation of Rho GTPase.

Here we discover that the FMF disease protein Pyrin is a specific immune sensor for bacterial modifications of Rho GTPases. Common to all identified bacterial stimuli is the modification of a switch-I residue in Rho-subfamily and GTPase inactivation. The modifications cover glucosylation, adenylylation, ADP-ribosylation and deamidation occurring on different residues. Thus, Pyrin does not directly recognize Rho modification, but probably senses an event downstream of Rho modification in the actin cytoskeleton pathway. Interestingly, direction interaction of Pyrin with actin and co-localization of Pyrin–ASC complex with actin filaments are observed[29]. Pyrin detects pathogen virulence activity, which is different from most mammalian PRRs that directly recognize microbial products. The disease resistance PRR protein in plant innate immunity often acts in an indirect manner by detecting pathogen-induced modification of a host protein, a model known as guard hypothesis[30]. Thus, the mode of Pyrin action echoes the plant guard model to some extent.

METHODS SUMMARY

Purified recombinant TcdB or TcsL was added into serum-free macrophage culture medium at a final concentration of 0.1 μg ml^{-1} or that indicated for 2.5–3 h. Supernatants of BMDMs (5×10^5) or DC2.4 cells (1×10^6) were subjected to trichloroacetic acid precipitation and the precipitates were analysed by anti-caspase 1 immunoblotting to detect pro-caspase 1 (p45) and the processed p10 fragment; cell lysates were blotted with anti-Actin antibody to ensure equal loading. To measure IL-1β secretion, BMDMs or DC2.4 cells were primed with LPS (500 ng ml^{-1}, 2 h), and released mature IL-1β was determined using the IL-1β ELISA kit (Neobioscience Technology Company). Cell pyroptosis was measured by the lactate dehydrogenase assay using CytoTox 96 Non-Radioactive Cytotoxicity Assay kit (Promega). All independent experiments carried out in this study and indicated in the figure legends were biological replicates.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 October 2013; accepted 7 May 2014.
Published online 11 June 2014.

1. Seshadri, S., Duncan, M. D., Hart, J. M., Gavrilin, M. A. & Wewers, M. D. Pyrin levels in human monocytes and monocyte-derived macrophages regulate IL-1β processing and release. *J. Immunol.* **179,** 1274–1281 (2007).
2. Yu, J. W. *et al.* Cryopyrin and pyrin activate caspase-1, but not NF-κB, via ASC oligomerization. *Cell Death Differ.* **13,** 236–249 (2006).
3. French FMF Consortium. A candidate gene for familial Mediterranean fever. *Nature Genet.* **17,** 25–31 (1997).
4. The International FMF Consortium. Ancient missense mutations in a new member of the *RoRet* gene family are likely to cause familial Mediterranean fever. *Cell* **90,** 797–807 (1997).
5. Chae, J. J. *et al.* Gain-of-function Pyrin mutations induce NLRP3 protein-independent interleukin-1β activation and severe autoinflammation in mice. *Immunity* **34,** 755–768 (2011).
6. Jank, T. & Aktories, K. Structure and mode of action of clostridial glucosylating toxins: the ABCD model. *Trends Microbiol.* **16,** 222–229 (2008).
7. Kuehne, S. A. *et al.* The role of toxin A and toxin B in *Clostridium difficile* infection. *Nature* **467,** 711–713 (2010).
8. Lyras, D. *et al.* Toxin B is essential for virulence of *Clostridium difficile*. *Nature* **458,** 1176–1179 (2009).
9. Kayagaki, N. *et al.* Non-canonical inflammasome activation targets caspase-11. *Nature* **479,** 117–121 (2011).
10. Ng, J. *et al. Clostridium difficile* toxin-induced inflammation and intestinal injury are mediated by the inflammasome. *Gastroenterology* **139,** 542–552.e1–3 (2010).
11. Keestra, A. M. *et al.* Manipulation of small Rho GTPases is a pathogen-induced process detected by NOD1. *Nature* **496,** 233–237 (2013).
12. Fukazawa, A. *et al.* GEF-H1 mediated control of NOD1 dependent NF-κB activation by *Shigella* effectors. *PLoS Pathog.* **4,** e1000228 (2008).
13. Richards, N. *et al.* Interaction between pyrin and the apoptotic speck protein (ASC) modulates ASC-induced apoptosis. *J. Biol. Chem.* **276,** 39320–39329 (2001).
14. Cordero, C. L., Kudryashov, D. S., Reisler, E. & Satchell, K. J. The actin cross-linking domain of the *Vibrio cholerae* RTX toxin directly catalyzes the covalent cross-linking of actin. *J. Biol. Chem.* **281,** 32366–32374 (2006).
15. Sheahan, K. L. & Satchell, K. J. Inactivation of small Rho GTPases by the multifunctional RTX toxin from *Vibrio cholerae*. *Cell. Microbiol.* **9,** 1324–1335 (2007).
16. Aktories, K. Bacterial protein toxins that modify host regulatory GTPases. *Nature Rev. Microbiol.* **9,** 487–498 (2011).
17. Yarbrough, M. L. *et al.* AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. *Science* **323,** 269–272 (2009).
18. Mattoo, S. *et al.* Comparative analysis of *Histophilus somni* immunoglobulin-binding protein A (IbpA) with other fic domain-containing enzymes reveals differences in substrate and nucleotide specificities. *J. Biol. Chem.* **286,** 32834–32842 (2011).
19. Worby, C. A. *et al.* The fic domain: regulation of cell signaling by adenylylation. *Mol. Cell* **34,** 93–103 (2009).
20. Popoff, M. R. *et al.* Ras, Rap, and Rac small GTP-binding proteins are targets for *Clostridium sordellii* lethal toxin glucosylation. *J. Biol. Chem.* **271,** 10217–10224 (1996).
21. Just, I., Selzer, J., Hofmann, F., Green, G. A. & Aktories, K. Inactivation of Ras by *Clostridium sordellii* lethal toxin-catalyzed glucosylation. *J. Biol. Chem.* **271,** 10149–10153 (1996).
22. Aktories, K. & Hall, A. Botulinum ADP-ribosyltransferase C_3: a new tool to study low molecular weight GTP-binding proteins. *Trends Pharmacol. Sci.* **10,** 415–418 (1989).
23. Jank, T., Giesemann, T. & Aktories, K. *Clostridium difficile* glucosyltransferase toxin B-essential amino acids for substrate binding. *J. Biol. Chem.* **282,** 35222–35231 (2007).
24. Boyer, L. *et al.* Pathogen-derived effectors trigger protective immunity via activation of the Rac2 enzyme and the IMD or Rip kinase signaling pathway. *Immunity* **35,** 536–549 (2011).
25. Bruno, V. M. *et al. Salmonella* Typhimurium type III secretion effectors stimulate innate immune responses in cultured epithelial cells. *PLoS Pathog.* **5,** e1000538 (2009).
26. Flannagan, R. S. *et al. Burkholderia cenocepacia* disrupts host cell actin cytoskeleton by inactivating Rac and Cdc42. *Cell. Microbiol.* **14,** 239–254 (2012).
27. Rosales-Reyes, R., Skeldon, A. M., Aubert, D. F. & Valvano, M. A. The Type VI secretion system of *Burkholderia cenocepacia* affects multiple Rho family GTPases disrupting the actin cytoskeleton and the assembly of NADPH oxidase complex in macrophages. *Cell. Microbiol.* **14,** 255–273 (2012).
28. Gavrilin, M. A. *et al.* Activation of the pyrin inflammasome by intracellular *Burkholderia cenocepacia*. *J. Immunol.* **188,** 3469–3477 (2012).
29. Waite, A. L. *et al.* Pyrin and ASC co-localize to cellular sites that are rich in polymerizing actin. *Exp. Biol. Med. (Maywood)* **234,** 40–52 (2009).
30. Jones, J. D. & Dangl, J. L. The plant immune system. *Nature* **444,** 323–329 (2006).

Acknowledgements We thank V. Dixit for providing knockout mice and anti-ASC antibody, D. Lyras for *C. sordellii* genomic DNA, J. Xiao for IbpA-Fic constructs, T. Iida for *V. parahaemolyticus* strain, M. Valvano for pDAI-SceI vector and H. Feng for TcdB *B. megaterium* expression system. We thank members of the Shao laboratory for discussions. The research was supported in part by an International Early Career Scientist grant from the Howard Hughes Medical Institute to F.S. This work was also supported by the National Basic Research Program of China 973 Program (2012CB518700), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB08020202) and China National Science Foundation Program for Distinguished Young Scholars (31225002) to F.S.

Author Contributions F.S. conceived the study; H.X. and J.Y. performed the majority of experiments, assisted by W.G.; L.L., P.L., L.Z., Y.-N.G., X.P., J.J.X., S.C. and F.W. contributed reagents and analytic tools. H.X., J.Y. and F.S. analysed the data and wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.S. (shaofeng@nibs.ac.cn).

LETTER

doi:10.1038/nature13459

Viral tagging reveals discrete populations in *Synechococcus* viral genome sequence space

Li Deng[1]†*, J. Cesar Ignacio-Espinoza[2]*, Ann C. Gregory[1], Bonnie T. Poulos[1], Joshua S. Weitz[3,4], Philip Hugenholtz[5] & Matthew B. Sullivan[1,2]

Microbes and their viruses drive myriad processes across ecosystems ranging from oceans and soils to bioreactors and humans[1–4]. Despite this importance, microbial diversity is only now being mapped at scales relevant to nature[5], while the viral diversity associated with any particular host remains little researched. Here we quantify host-associated viral diversity using viral-tagged metagenomics, which links viruses to specific host cells for high-throughput screening and sequencing. In a single experiment, we screened 10^7 Pacific Ocean viruses against a single strain of *Synechococcus* and found that naturally occurring cyanophage genome sequence space is statistically clustered into discrete populations. These population-based, host-linked viral ecological data suggest that, for this single host and seawater sample alone, there are at least 26 double-stranded DNA viral populations with estimated relative abundances ranging from 0.06 to 18.2%. These populations include previously cultivated cyanophage and new viral types missed by decades of isolate-based studies. Nucleotide identities of homologous genes mostly varied by less than 1% within populations, even in hypervariable genome regions, and by 42–71% between populations, which provides benchmarks for viral metagenomics and genome-based viral species definitions. Together these findings showcase a new approach to viral ecology that quantitatively links objectively defined environmental viral populations, and their genomes, to their hosts.

Decades-old microscopic observations revealed that viruses typically outnumber microbial cells approximately tenfold in marine systems[1], recasting them from environmentally insignificant to the most abundant biological entities on Earth. Viruses are now considered important in microbial mortality, horizontal gene transfer and global biogeochemistry[2,3], with recent recognition of vast cellular metabolic reprogramming[4] during infection. However, the enormous microbial and viral diversity in nature makes it challenging to clarify and quantify these roles, particularly as viral taxonomy remains largely based on morphology and properties of isolates. Although large-scale isolate-based sequencing studies are clarifying genomic parameters for viral taxonomy—for example, defining phage 'genus' boundaries[6–8]—they remain limited to cultivated viral groups that represent only a fraction of viruses in nature.

Here we explore genetic variation in an environmentally relevant cyanobacterial model system[5,9,10]—seawater cyanophages within a Pacific Ocean viral assemblage that infect a cultured cyanobacterial host. We do so by adapting viral tagging, a high-throughput means of linking viruses to a target host[11], for use in the field. In this method, DNA in environmental viruses is labelled non-specifically with a fluorescent dye, viruses are mixed with a 'bait host' pre-labelled with isotopically heavy DNA, and infected cells are collected by fluorescence-activated flow cytometry. Isotopically light viral DNA is then separated from heavy host DNA using a density gradient, and the infecting viral DNA is quantitatively amplified[12] to produce viral-tagged metagenomes. Beyond the identification of viral populations interacting with a particular host[13], the data shed light on lineage-specific viral ecology at scales not previously possible, enabling the development of population-based measurements and models of viral ecology and evolution.

To explore Pacific Ocean cyanophage diversity linked to the cyanobacterial host *Synechococcus* sp. WH7803, we applied a traditional culture-based approach complemented by metagenomic analysis of the double-stranded (ds)DNA viral community and viral-tagged community. Of 97 new isolates, 90 were myoviruses as inferred using a marker gene (Extended Data Fig. 1), which is consistent with previous isolates on this host (88% are myoviruses; Methods). Similarly, metagenomic analysis showed that viral tagging simplified the total viral community towards one dominated by myoviruses (Extended Data Fig. 2 and Supplementary Data 1). Viral tagging an artificial viral assemblage did not enrich for myoviruses (Methods), indicating that the *Synechococcus* WH7803–myovirus interactions are specific. Furthermore, these viruses are likely to infect, rather than just adsorb to, their host given previous and current experiments in which all tested cyanophage–host interactions led to infection when positively viral tagged (5 of 5 isolates tested previously[11], and 18 of 18 isolates tested here; Extended Data Table 1) .

Beyond the expected myoviruses, viral tagging also provided evidence (genomic data) for 42 new uncultured viruses specific to *Synechococcus* WH7803 (Extended Data Fig. 3 and Supplementary Data 1), including eight podoviruses (T7-like, phiKMV-like) and one siphovirus, as well as 33 partial genomes (contigs) that were ambiguous or lacked similarity to any known viral or bacterial genes, which may represent novel viruses (Methods). The screening of $\sim 10^7$ virus particles against *Synechococcus* WH7803 probably explains why such an unprecedented diversity of specific viruses were recovered for this single host despite two decades of isolation studies.

Viral-tagging-based screening of the bulk viral community improved assembly (average contig size increased from 1.2 kilobases (kb) to 5.5 kb) to produce three nearly complete genomes (*Candidatus* genomes; CG-01, CG-03 and CG-05; 197 kb, 185 kb and 108 kb, respectively, contigs containing 94–97% of 65 T4-like core genes; Table 1) and 164 viral contigs (Supplementary Data 1) that offer genomic context and enable host-specific discoveries. Auxiliary metabolic genes[14] previously observed in viral metagenomes can now be assigned to a discrete viral entity with an experimentally defined host. For example, membrane protein 'T17', antioxidant protein 'T768' and glycosyltransferase 'T1338' were assigned to T4-like cyanophages (Supplementary File 1). Conversely, the deep sampling of *Synechococcus* cyanophages did not identify any photosystem I (PSI) genes reported in putative cyanophage metagenomic fragments[15] but lacking in cyanophage genomes[9,10,16,17], suggesting that viral-encoded PSI genes are restricted to particular locations and/or hosts.

Insights from host-linked viral-tagging data also address a fundamental and persistent challenge in microbial ecology and evolution: how to define populations and thus quantify natural community diversity. Previous marine work suggests that genomes of co-existing viruses, infecting a

[1]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85719, USA. [2]Department of Molecular and Cellular Biology, University of Arizona, Tucson, Arizona 85719, USA. [3]School of Biology, Georgia Institute of Technology, Atlanta, Georgia 30332, USA. [4]School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, USA. [5]Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences & Institute for Molecular Bioscience, The University of Queensland, St Lucia QLB 4072, Australia. †Present address: Helmholtz Zentrum München-German Research Center for Environmental Health, Institute of Groundwater Ecology, Neuherberg 85764, Germany.
*These authors contributed equally to this work.

Table 1 | Relative abundance of T4-like myovirus populations in Pacific Ocean sea water

Rank	Name	Size (kb)	Percentage finished	No. of reads mapped (× 1,000)	Percentage of viral-tagged metagenome	Mean coverage	Percentage of isolates (PCR)
1	CG-05	108	59	2,313	18.20	1,019.0	22
	S-MbCM6/25*	176	NA				
2	CG-24	43	8	329	2.58	964.4	ND
3	CG-07	67	40	488	3.84	916.0	ND
4	CG-03	185	95	1,016	7.99	687.6	ND
5	CG-01	197	94	1,038	8.16	658.4	ND
6	CG-02	180	97	656	5.16	456.6	7.3
7	CG-09	83	57	201	1.58	304.3	ND
8	CG-06	117	37	149	1.18	159.1	ND
9	S-MbC100*	170	NA	127	1.00	93.4	17.1
10	CG-11	37	39	133	1.05	72.1	7.3
	S-MbCM7*	189	NA				
11	CG-04	114	36	65	0.51	71.4	ND
12	CG-19	44	6	25	0.20	69.8	ND
13	CG-12	33	5	18	0.14	69.6	ND
14	CG-13	42	6	21	0.17	63.9	ND
15	CG-10	68	11	29	0.23	53.1	12.2
16	CG-25	40	25	8	0.06	24.4	ND

Percentage finished refers to the estimated percentage of the complete genome captured, calculated as the fraction of the 65-gene T4 core genome observed in the resulting contig. Percentage of viral-tagging metagenome refers to the fraction of the viral-tagging metagenome reads present in the *Candidatus* genome or isolate genomes. Mean coverage refers to the average depth of coverage per *Candidatus* genome. Percentage of isolates (PCR) refers to the percentage of 41 isolates for which g23 sequences can be mapped to the *Candidatus* genomes or isolate genomes (identity >95%, only 41 of 97 g23 products of isolates were sequenced). CG, *Candidatus* genome (phylogenetically informative contig larger than 30 kb derived from the viral-tagging experimental data); NA, not applicable; ND, not detected; PCR, polymerase chain reaction.
* Isolates.

single host, range from relatively dissimilar (two co-isolated cyanophages[9] shared four-fifths of their genes at ~83% average amino acid identity (AAI)) to nearly identical (five roseophages[18] with ~97% average nucleotide identity (ANI)[19]). Genome sequences of hundreds of mycobacteriophages isolated using a single host have revealed 'rampant mosaicism' such that individual viral genomes are composed of assemblages of modules that challenge notions of demarcated populations and hierarchical, genome-based taxonomy (for example, see refs 20–23). Nonetheless, the mycobacteriophage sequences can be clustered into groups by nucleotide similarity, with within-group ANIs ranging from 63–99% (refs 6, 22). As in the marine case, whether these groups denote viral 'species' (that is, discrete ecological and evolutionary units) cannot be discerned given such broad ANI ranges and with only one or two phages sampled per site. Isolate-based genomics could be informative if scaled up, but a single viral-tagging experiment provides the opportunity, now, to explore this question and make four key inferences from its first field application.

First, dsDNA cyanophage genome sequence space is not a genetic continuum in nature, at least not for this particular phage type, host and site. Here, genome-wide genetic relatedness proxies[19] from conserved regions of the dominant T4-like cyanophages (Extended Data Fig. 4) generated a 'population genome landscape', revealing statistically significant discrete clustering of the viral-tagging sequences (Fig. 1) that are robust to variations in recruitment parameters (Extended Data Fig. 5). Such clustering is consistent with globally sampled mycobacteriophage groups[6,22] (see earlier), as well as population structure inferred in cyanophages using marker genes at single sites[24] and globally sampled genomes[25] (Extended Data Fig. 6), and in single-stranded (ss)DNA phages using genomes assembled from pooled natural samples interrogated by feature frequent profile analysis[26]. Yet, the viral-tagging data expand these findings by large-scale analysis at a single site to reveal discrete dsDNA viral clusters—that is, non-overlapping 'clouds' of metagenome-derived cyanophage sequence space—herein termed 'populations'. Whether these

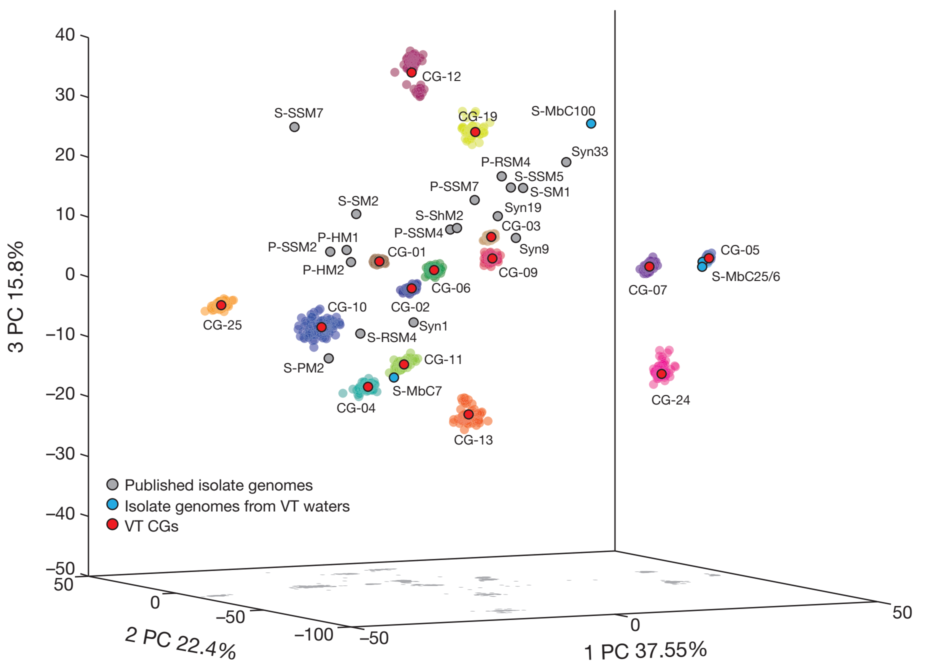


Figure 1 | Population genome landscape plot showing the genetic relationship of cultivated and viral-tagged T4-like phages of *Synechococcus* WH7803 from a single seawater sample and all available marine cyanophage genomes. Principal component (PC) projection of population-level variation within randomly resampled metagenomic data for each *Candidatus* genome (CG) recovered from the viral-tagged (VT) metagenome (the coloured 'clouds'; cloud colours are arbitrary to aid in discriminating populations). *Candidatus* genomes (phylogenetically informative contig, larger than 30 kb) were derived from the viral-tagging metagenome; accuracy of read assignment $Q = 0.9926$, Z-score = 142.2; Dunn index = 0.26, Z-score = 1,829.

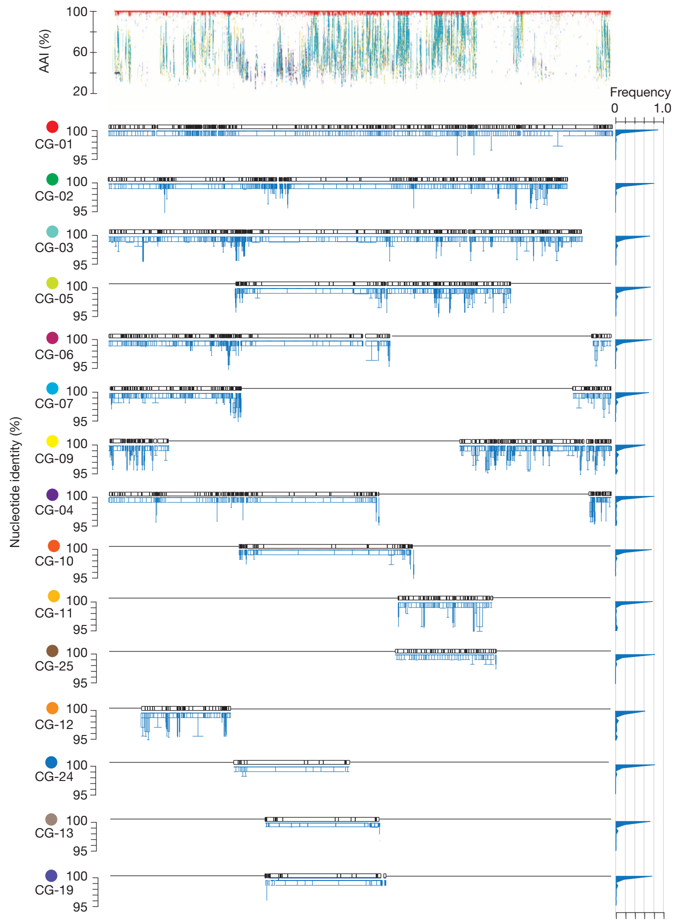


Figure 2 | The 15 dominant T4-like *Candidatus* genomes assembled from the viral-tagged metagenome. Top, reads that map to the *Candidatus* genome CG-01 (genome size 197 kb) by commonly used fragment recruitment metrics (BLASTx *e*-value < 0.001); dots represent reads assigned to the *Candidatus* genomes shown at the bottom, the match is indicated by the colour. Bottom, alignments of all T4-like *Candidatus* genomes against CG-01 are shown. The locus-by-locus nucleotide divergence of each open reading frame (blue) are plotted underneath each genome (0.09, 0.91, second and third quartile and median are shown). The histograms on the right show the summed genome-wide locus-to-locus variation. Note that most variation is concentrated in the top 1%.

populations formally represent species or not will require whole genome information and consideration of neutral and adaptive processes shaping observed variation[27].

Second, such discrete populations enable host-linked, population-based viral ecology. In this single seawater sample and for this host, there are at least 26 viral populations (a 27th is added through isolations, see later). These include 15 T4-like phage populations (Table 1), three of which include co-isolated genomes (Extended Data Fig. 7a), as well as 11 non-T4-like phage populations (Extended Data Fig. 3c). This estimate of 15 T4-like phage populations is consistent with maximal coverage depth in the larger data set of T4 contigs collected here (Extended Data Fig. 7b). Together, these 26 populations represented 0.05–18.2% of the viral-tagging metagenome reads (Table 1 and Extended Data Fig. 3c), with ~53% of the reads assignable to these populations and up to 60% if all small contigs are considered. The remaining 40% of the viral-tagging metagenome reads probably represent reads from the 'rare virosphere' (Methods). The per population, metagenome-derived coverage serves as a proxy for abundance, which enables an estimate of the first host-associated wild viral rank–abundance distribution (Extended Data Fig. 7c), analogous to long-standing ecological efforts to characterize species abundances in natural communities[28].

Third, viral tagging allows quantitative examination of cyanophage culture bias. Here, four *Myoviridae* isolates from the same waters included the first, ninth and tenth most abundant T4-like phage populations observed in the viral-tagging metagenomes on this host (Table 1). In addition, all amplicons derived from isolates using PCR with primers that target the major capsid protein (gp23) from T4-like myoviruses can be mapped to the viral-tagging populations (first, sixth, ninth, tenth, fifteenth in Table 1, and the rest to the small contigs, see Methods). This overlap between isolates and viral-tagging populations partially validates the viral-tagging procedure and suggests that, at least for T4-like cyanophages, culture bias might be relatively minimal. Notably, however, no isolates showed similarity to the 42 new viruses revealed by viral tagging, and the projected variation in sequence space recovered by a single viral-tagging experiment is larger than that associated with published global isolates (Fig. 1). Together, this suggests that culture-based studies may miss major routes of horizontal gene transfer and/or ecological interactions.

Last, we were able to document intra-population variation for wild uncultured viruses (Fig. 2), which is critical for interpreting metagenomic fragment recruitment analyses and establishing a genome-based viral species definition. Here, each population's locus-to-locus, pairwise percent nucleotide identity between the reference sequence and its 'assigned' viral-tagging metagenomic reads ranged from 95–100% ANI (mean 99.53%; for example, see insets in Fig. 2), with some populations varying more than others (see spread of clouds in Fig. 1 and box plots in Fig. 2). This is similar to >99% ANI observed across eight loci used to group 60 isolates into five clusters in *Synechococcus* cyanophage isolates[24], and >95% ANI commonly associated with microbial species definitions[19]. However, it is more conservative than most of the range (83–97%, average 90%) of ANIs observed in ten isolates in the phiKMV species complex[29]. Interestingly, intra-population ANIs from conserved and hypervariable regions are statistically indistinguishable (Fig. 2). By contrast, pairwise inter-population variation observed in the viral-tagging metagenomes suggests that nucleotide identities range from 42 to 71% between populations (Extended Data Table 2). The finding of high intra-population ANI from hypervariable regions of the captured cyanophage stands in contrast to models of rampant phage mosaicism[22], in which assemblages of modules within viral genomes suggest a horizontal, rather than vertical, evolutionary signal. It remains to be determined if this observation is exceptional or the rule for phage population structure. Similar intra- and inter-population sequence divergence levels are maintained by differences in relative recombination rates in bacteria and archaea[27]. Formal testing of whole genome data in a population genetic framework (for example, see ref. 27) is needed to assess the validity of these empirically derived populations as species. Nonetheless, these viral-tagging data already provide a much-needed benchmark for refining metagenomic analyses, albeit from a single host and sample, by suggesting an empirical cut-off (<95% ANI) for reads that probably derive from different populations.

In conclusion, viral-tagging-enabled experimental linkage of wild cyanophages to their host at a single site provides evidence that phage genome sequence space is structured in nature, just as recently posited for bacteria and archaea[27,30]. Moving forward, viral tagging has the potential to enable researchers to broadly map how viruses change over space and time. Given that such comparative viral-tagging data are genome- and host-linked, as well as population-based, these data should better elucidate the processes that drive viral population structure in nature.

METHODS SUMMARY

Viral-tagging metagenomes. Surface (10 m) water was collected from Station H3 (36° 73.5 N, 237° 98.1 E) in Monterey Bay, California, United States, 0.22 μm filtered and stored (4 °C, dark). *Synechococcus* WH7803, labelled with ^{15}N, was mixed with fluorescently labelled viral particles[11], then viral-tagged cells (increased fluorescence) were flow-cytometrically sorted, DNA was extracted, and ^{15}N-labelled heavy host DNA was separated from non-labelled light viral DNA by CsCl density ultracentrifugation. In total, 3×10^8 virions were co-incubated with 3×10^7 host cells for

60 min then 1.2×10^7 viral-tagged cells were sorted and used for DNA extraction. Light DNA was linker amplified[12] for sequencing.

Community metagenomes. Viral concentrates were prepared from 20 l of 0.22 μm filtrate by chemical flocculation and purified using ultracentrifugation. DNA was extracted, linker amplified[12] and sequenced. However, our metagenomic DNA preparation method would strongly select against ssDNA phages, and not capture RNA phages at all.

Phage isolation and characterization. Ninety-seven cyanophages able to infect *Synechococcus* WH7803 were isolated and purified as previously described[16]. Ninety of 97 isolates were assigned to T4-like myoviruses using a specific gene marker (gp20). Four isolate genomes were assembled completely: S-MbCM6, S-MbCM7, S-MbCM25 and S-MbCM100.

Bioinformatic analyses. Quality control, filtering, assembly, protein clustering, annotation, taxonomy analyses, whole genome comparison, statistics, locus-by-locus variation and associated bioinformatics analyses were done using a set of custom scripts detailed in Methods and Extended Data Fig. 8. Scripts and associated documentation are available at http://www.eebweb.arizona.edu/faculty/mbsulli/informatics.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 5 December 2013; accepted 7 May 2014.
Published online 13 July 2014.

1. Bergh, O., Børsheim, K. Y., Bratbak, G. & Heldal, M. High abundance of viruses found in aquatic environments. *Nature* **340,** 467–468 (1989).
2. Breitbart, M. Marine viruses: truth or dare. *Annu. Rev. Mar. Sci.* **4,** 425–448 (2012).
3. Suttle, C. A. Marine viruses—major players in the global ecosystem. *Nature Rev. Microbiol.* **5,** 801–812 (2007).
4. Hurwitz, B. L., Hallam, S. J. & Sullivan, M. B. Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biol.* **14,** R123 (2013).
5. Flombaum, P. *et al.* Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc. Natl Acad. Sci. USA* **110,** 9824–9829 (2013).
6. Holmfeldt, K., Odić, D., Sullivan, M. B., Middelboe, M. & Riemann, L. Cultivated single-stranded DNA phages that infect marine Bacteroidetes prove difficult to detect with DNA-binding stains. *Appl. Environ. Microbiol.* **78,** 892–894 (2012).
7. Hatfull, G. F., Jacobs-Sera, D. & Lawrence, J. G. Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. *J. Mol. Biol.* **397,** 119–143 (2010).
8. Lavigne, R., Seto, D., Mahadevan, P., Ackermann, H.-W. & Kropinski, A. M. Unifying classical and molecular taxonomic classification: analysis of the *Podoviridae* using BLASTP-based tools. *Res. Microbiol.* **159,** 406–414 (2008).
9. Sullivan, M. B. *et al.* Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. *Environ. Microbiol.* **12,** 3035–3056 (2010).
10. Mann, N. H. *et al.* The genome of S-PM2, a 'photosynthetic' T4-type bacteriophage that infects marine *Synechococcus* strains. *J. Bacteriol.* **187,** 3188–3200 (2005).
11. Deng, L., Gregory, A., Yilmaz, S., Poulos, B. T., Hugenholtz, P. & Sullivan, M. B. Contrasting life strategies of viruses that infect photo- and heterotrophic bacteria, as revealed by viral tagging. *MBio* **3,** e00373–12 (2012).
12. Duhaime, M. B., Deng, L., Poulos, B. T. & Sullivan, M. B. Towards quantitative metagenomics of wild viruses and other ultra-low concentration DNA samples: a rigorous assessment and optimization of the linker amplification method. *Environ. Microbiol.* **14,** 2526–2537 (2012).
13. Weitz, J. S. *et al.* Phage–bacteria infection networks. *Trends Microbiol.* **21,** 82–91 (2013).
14. Sharon, I. *et al.* Comparative metagenomics of microbial traits within oceanic viral communities. *ISME J.* **5,** 1178–1190 (2011).
15. Sharon, I. *et al.* Photosystem I gene cassettes are present in marine virus genomes. *Nature* **461,** 258–262 (2009).
16. Sullivan, M. B., Coleman, M. L., Weigele, P., Rohwer, F. & Chisholm, S. W. Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. *PLoS Biol.* **3,** e144 (2005).
17. Millard, A. D., Zwirglmaier, K., Downey, M. J., Mann, N. H. & Scanlan, D. J. Comparative genomics of marine cyanomyoviruses reveals the widespread occurrence of *Synechococcus* host genes localized to a hyperplastic region: implications for mechanisms of cyanophage evolution. *Environ. Microbiol.* **11,** 2370–2387 (2009).
18. Angly, F. *et al.* Genomic analysis of multiple Roseophage SIO1 strains. *Environ. Microbiol.* **11,** 2863–2873 (2009).
19. Konstantinidis, K. T. Genomic insights that advance the species definition for prokaryotes. *Proc. Natl Acad. Sci. USA* **102,** 2567–2572 (2005).
20. Lawrence, J. G., Hatfull, G. F. & Hendrix, R. W. Imbroglios of viral taxonomy: genetic exchange and failings of phenetic approaches. *J. Bacteriol.* **184,** 4891–4905 (2002).
21. Hendrix, R. W., Lawrence, J. G., Hatfull, G. F. & Casjens, S. The origins and ongoing evolution of viruses. *Trends Microbiol.* **8,** 504–508 (2000).
22. Hatfull, G. F. The secret lives of mycobacteriophages. *Adv. Virus Res.* **82,** 179–288 (2012).
23. Hendrix, R. W., Smith, M. C., Burns, R. N., Ford, M. E. & Hatfull, G. F. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc. Natl Acad. Sci. USA* **96,** 2192–2197 (1999).
24. Marston, M. F. & Amrich, C. G. Recombination and microdiversity in coastal marine cyanophages. *Environ. Microbiol.* **11,** 2893–2903 (2009).
25. Ignacio-Espinoza, J. C. & Sullivan, M. B. Phylogenomics of T4 cyanophages: lateral gene transfer in the 'core' and origins of host genes. *Environ. Microbiol.* **14,** 2113–2126 (2012).
26. Labonté, J. M. & Suttle, C. A. Previously unknown and highly divergent ssDNA viruses populate the oceans. *ISME J.* **7,** 2169–2177 (2013).
27. Polz, M. F., Alm, E. J. & Hanage, W. P. Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends Genet.* **29,** 170–175 (2013).
28. Whittaker, R. H. Dominance and diversity in land plant communities: numerical relations of species express the importance of competition in community function and evolution. *Science* **147,** 250–260 (1965).
29. Ceyssens, P. J. *et al.* Phenotypic and genotypic variations within a single bacteriophage species. *Virol. J.* **8,** 134 (2011).
30. Shapiro, B. J. *et al.* Population genomics of early events in the ecological differentiation of bacteria. *Science* **336,** 48–51 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements Funding was provided by the US Department of Energy (DOE) Joint Genome Institute (JGI) Community Sequencing Program, Biosphere 2, BIO5, US National Science Foundation (NSF) OCE0940390, and Gordon and Betty Moore Foundation grants to M.B.S., as well as NSF OCE1233760 and Burroughs Wellcome Fund grants to J.S.W. We thank J. Fuhrman for suggesting stable-isotope-labelled host DNA; A. Z. Worden and the CANON Initiative for the cruise opportunity; Worden laboratory members; the captain and crew of the R/V Western Flyer for operational/sampling support; J. B. Waterbury, S. W. Chisholm and A. Wichels for strains; the Tucson Marine Phage laboratory; Institute of Groundwater Ecology of Helmholtz Munich; and N. Pace, M. Young, S. W. Chisholm and S. Yilmaz for technical/analytical support and manuscript comments. We acknowledge the University of Arizona Genetics Core for viral-tagging metagenomic sequencing; iCyt and the Arizona Cancer Center and Arizona Research Laboratories (ARL) Division of Biotechnology Cytometry Core Facility for cytometry support; the University Information Technology Services Research Computing Group and the ARL Biotechnology Computing for high-performance computing clusters (HPCC) access and support. Community metagenomic sequencing was provided by the DOE JGI Community Sequencing Program under the Office of Science of the US DOE contract no. DE-AC02-05CH11231.

Author Contributions L.D., P.H. and M.B.S. designed the experiments. L.D. collected samples. L.D., A.C.G. and B.T.P. performed the experiments. L.D., J.C.I.-E., J.S.W., P.H. and M.B.S. analysed data, interpreted results and wrote the paper.

Author Information Data for viral genomes have been deposited in GenBank under accession numbers JN371768 and KF156338-40; metagenomic data have been deposited in CAMERA under accession numbers CAM_P_0001068 and CAM_P_0000915; raw data including gp23 sequences and informatic pipelines, assemblies and data for figures are available at http://datadryad.org/resource/doi:10.5061/dryad.gr3ks. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.B.S. (mbsulli@email.arizona.edu).

doi:10.1038/nature13452

Carbonic anhydrases, *EPF2* and a novel protease mediate CO_2 control of stomatal development

Cawas B. Engineer[1], Majid Ghassemian[2], Jeffrey C. Anderson[3], Scott C. Peck[3], Honghong Hu[1]† & Julian I. Schroeder[1]

Environmental stimuli, including elevated carbon dioxide levels, regulate stomatal development[1–3]; however, the key mechanisms mediating the perception and relay of the CO_2 signal to the stomatal development machinery remain elusive. To adapt CO_2 intake to water loss, plants regulate the development of stomatal gas exchange pores in the aerial epidermis. A diverse range of plant species show a decrease in stomatal density in response to the continuing rise in atmospheric CO_2 (ref. 4). To date, one mutant that exhibits deregulation of this CO_2-controlled stomatal development response, *hic* (which is defective in cell-wall wax biosynthesis, ref. 5), has been identified. Here we show that recently isolated *Arabidopsis thaliana* β-carbonic anhydrase double mutants (*ca1 ca4*)[6] exhibit an inversion in their response to elevated CO_2, showing increased stomatal development at elevated CO_2 levels. We characterized the mechanisms mediating this response and identified an extracellular signalling pathway involved in the regulation of CO_2-controlled stomatal development by carbonic anhydrases. RNA-seq analyses of transcripts show that the extracellular pro-peptide-encoding gene *EPIDERMAL PATTERNING FACTOR 2* (*EPF2*)[7,8], but not *EPF1* (ref. 9), is induced in wild-type leaves but not in *ca1 ca4* mutant leaves at elevated CO_2 levels. Moreover, *EPF2* is essential for CO_2 control of stomatal development. Using cell-wall proteomic analyses and CO_2-dependent transcriptomic analyses, we identified a novel CO_2-induced extracellular protease, CRSP (CO_2 RESPONSE SECRETED PROTEASE), as a mediator of CO_2-controlled stomatal development. Our results identify mechanisms and genes that function in the repression of stomatal development in leaves during atmospheric CO_2 elevation, including the carbonic-anhydrase-encoding genes *CA1* and *CA4* and the secreted protease CRSP, which cleaves the pro-peptide EPF2, in turn repressing stomatal development. Elucidation of these mechanisms advances the understanding of how plants perceive and relay the elevated CO_2 signal and provides a framework to guide future research into how environmental challenges can modulate gas exchange in plants.

CO_2 exchange between plants and the atmosphere, and water loss from plants to the atmosphere, depends on the density and the aperture size of plant stomata, and plants have evolved sophisticated mechanisms to control this flux[1–3,10,11]. Ecophysiological studies have highlighted the importance of stomatal density in the context of global ecology and climate change[12]. Plants adapt to the continuing rise in atmospheric CO_2 concentration by reducing their stomatal density[4] (that is, the number of stomata per unit of epidermal surface area). This change causes the leaf temperature to rise because of a decrease in the plant's evapotranspirative cooling ability, while simultaneously increasing the transpiration efficiency of plants[13]. These phenomena, combined with the increasing scarcity of fresh water for agriculture, are predicted to dramatically impact on plant health[12,14,15].

In recent research, we identified mutations in the *A. thaliana* β-carbonic anhydrase genes *CA1* (At3g01500) and *CA4* (At1g70410) that impair the rapid, short-term CO_2-induced stomatal movement response[6]. Although *ca1 ca4* (double mutant) plants show a higher stomatal density than wild-type plants, it remains unknown whether CO_2 control of stomatal development is affected in these plants[6]. We investigated whether the long-term CO_2 control of stomatal development is altered in *ca1 ca4* plants. We analysed the stomatal index of wild-type (WT) and *ca1 ca4* plants grown at low (150 p.p.m.) and elevated (500 p.p.m.) CO_2 concentrations. For WT plants (Columbia (Col)), growth at the elevated CO_2 concentration resulted in, on average, 8% fewer stomata than growth at the low CO_2 concentration (Fig. 1a–c and Extended Data Fig. 1). The *ca1 ca4* mutant did not show an elevated CO_2-induced repression of the stomatal index; however, interestingly, *ca1 ca4* plants grown at the elevated CO_2 concentration showed an average 22% increase in the stomatal index in their cotyledons ($P < 0.024$; Fig. 1b, c) compared with *ca1 ca4* plants grown at the low CO_2 concentration. Similar results were obtained when stomatal density measurements were analysed (Fig. 1d). The mature rosette leaf phenotype in *ca1 ca4* mutants also showed an increase in the stomatal index at the elevated CO_2 concentration, which is consistent with the observations in the cotyledons (Extended Data Fig. 1a; stomatal indices rather than densities were analysed for accuracy; see Methods and Extended Data Fig. 1c legend).

We transformed the *ca1 ca4* mutant with genomic constructs expressing either *CA1* or *CA4* and investigated complementation of their stomatal development responses to CO_2. Five of six independent transformant lines for either the *CA1* or *CA4* gene showed a significant suppression of the elevated CO_2-induced inversion in the stomatal index found in *ca1 ca4* plants (Fig. 1e, f). By contrast, *ca1 ca4* leaves showed an average of 20% more stomata than WT leaves at the elevated CO_2 concentration. The complementation lines showed varying levels of suppression of the inverted stomatal development phenotype of *ca1 ca4* plants (Fig. 1e, f).

We tested the effects of preferential expression of these native *A. thaliana* carbonic anhydrases in mature guard cells[6,16], as yellow fluorescent protein (YFP) fusion proteins (Extended Data Fig. 2a–c). These cell-type-specific complementation analyses showed that the enhanced stomatal development in *ca1 ca4* plants at the elevated CO_2 concentration can be suppressed by preferential expression of either *CA1* or *CA4* in mature guard cells (Extended Data Fig. 2b–d). This result provides initial evidence for extracellular signalling in the CO_2 response mediated by these carbonic anhydrases during protodermal cell fate specification in developing cotyledons. It also indicates that the catalytic activity of the carbonic anhydrases may be required for CO_2 control of stomatal development (see Extended Data Fig. 1d for data on complementation analyses with an unrelated, human, carbonic anhydrase, CA-II). We note that although we can complement the *ca1 ca4* mutant phenotype with mature-guard-cell-targeted carbonic anhydrase overexpression, this finding does not exclude the possibility that expression in other cell types could function in this process. For example, in addition to being highly expressed in mature guard cells, *CA1* and *CA4* are also highly expressed in meristemoids, pavement cells and mesophyll cells[6,16,17]. Experiments analysing CO_2 control of stomatal development in the open stomata 1 mutant *ost1-3* show a divergence in the CO_2-mediated

[1]Division of Biological Sciences, University of California San Diego, La Jolla, California 92093, USA. [2]Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California 92093, USA. [3]Department of Biochemistry, University of Missouri-Columbia, Columbia, Missouri 65211, USA. †Present address: College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, China.

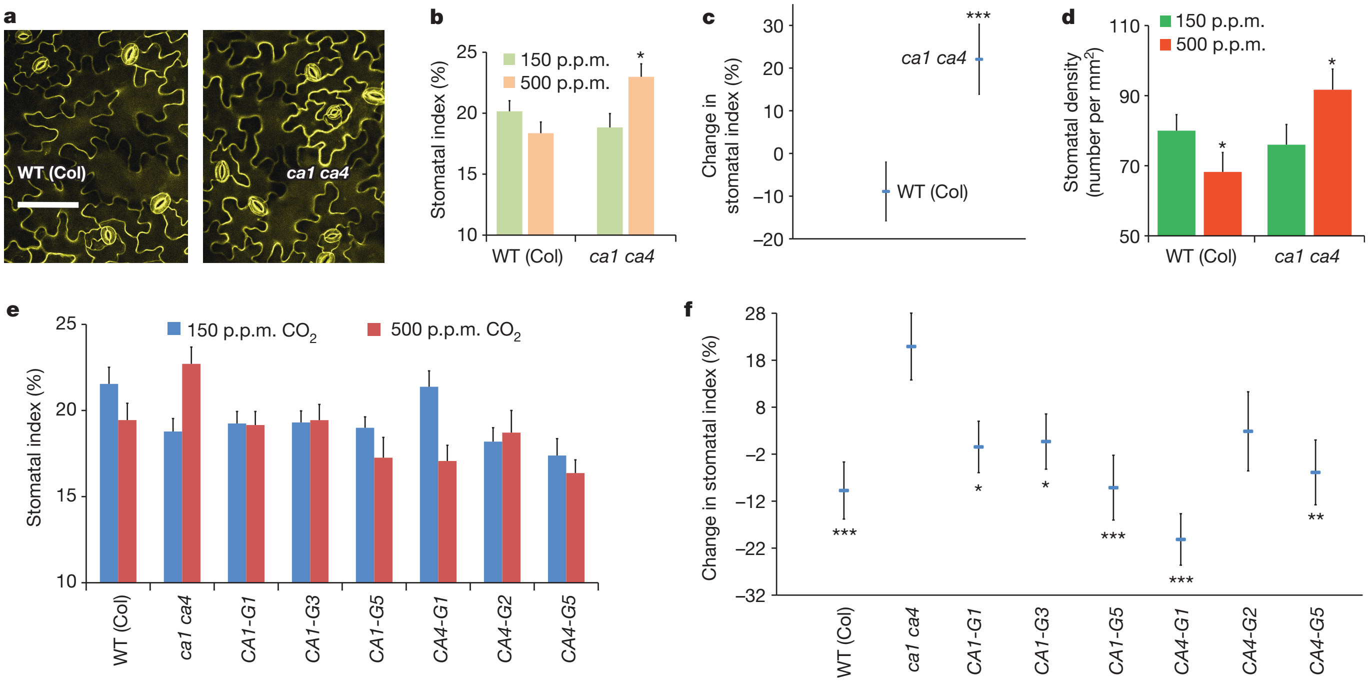


Figure 1 | The carbonic anhydrases *CA1* and *CA4* are required for repression of stomatal development at elevated CO_2 concentrations. **a**, Confocal images of the abaxial cotyledon epidermis of 10-day-old *ca1 ca4* and WT (Col) seedlings grown at 500 p.p.m. CO_2. Scale bar, 100 µm. **b**, Stomatal index of WT and *ca1 ca4* seedlings grown at 150 and 500 p.p.m. CO_2, showing an inverted stomatal development response to elevated CO_2 by the mutant. **c**, Elevated CO_2-induced changes in the stomatal index (data from **b**) shown as percentage changes in the stomatal index at 500 p.p.m. CO_2 relative to 150 p.p.m. CO_2. **d**, Stomatal density (data from **c**) for WT and *ca1 ca4* seedlings. **e**, Stomatal index for six independent complementation lines of *ca1 ca4* transformed with genomic copies of either *A. thaliana CA1* (*CA1-G*) or *A. thaliana CA4* (*CA4-G*). **f**, Elevated CO_2-induced changes in stomatal development (data from **e**). **b**–**f**, Statistical comparisons were made between CO_2 treatments (**b** and **d**) or were compared with the WT (**c**) or the *ca1 ca4* data (**f**). Stomatal density and index measurements were conducted on 10-day-old seedlings. Error bars show mean ± s.e.m., $n = 20$ for **b**–**f**. ***, $P < 0.00005$; **, $P < 0.005$; *, $P < 0.05$, using analysis of variance (ANOVA) and Tukey's post-hoc test.

signalling pathways controlling stomatal movements[18] and stomatal development (Extended Data Fig. 1e).

To gain initial insight into the regulatory mechanisms by which signalling in response to an elevated CO_2 concentration exerts CA1- and CA4-dependent repression of stomatal development, we conducted high-throughput RNA-seq transcriptomics on immature aerial tissues of *A. thaliana* seedlings grown at the low and elevated CO_2 concentrations. These analyses and independent single gene quantitative PCR (qPCR) studies of developing cotyledons showed that elevated CO_2 induced upregulation of transcripts of *EPF2* (which encodes an extracellular pro-peptide ligand)[7,8] in WT plants but not *ca1 ca4* plants (Fig. 2a). Our mature guard cell complementation analyses support a role for extracellular signalling in the elevated CO_2-mediated repression of stomatal development (Extended Data Figs 1d and 2).

EPF2 is an early mediator of protodermal cell fate specification and controls cell entry to the stomatal lineage by limiting asymmetric divisions[7,8]. *MUTE*[19,20] expression is a reliable indicator of cells that are committed to the stomatal lineage[19,20]. We transformed and examined WT and *ca1 ca4* plants harbouring a *MUTEpro*::*nucGFP* construct[19] (which allows expression of green fluorescent protein localized to the nucleus). Compared with WT plants, *ca1 ca4* plants expressed *MUTEpro*::*nucGFP* in 33% more cells, on average, at the elevated CO_2 concentration but not the low CO_2 concentration (Fig. 2b, c). The *MUTEpro*::*nucGFP* expression data provide an independent measure of the effect of *ca1 ca4* on the CO_2 response and are correlated with the increased stomatal index of *ca1 ca4* leaves that is found at the elevated CO_2 concentration (Fig. 1b). These data suggest that the increased stomatal development in *ca1 ca4* plants at the elevated CO_2 concentration progresses via components upstream of MUTE.

We analysed whether genetic perturbation of *EPF2* results in an abnormal stomatal development response to CO_2 concentration. Remarkably, plants carrying either of two independent mutant *epf2* alleles showed a clear inversion in CO_2 control of stomatal development (Fig. 2d and Extended Data Fig. 1b), with an average of 23% more stomata at the elevated CO_2 concentration than at the low concentration. We also tested the effects of a very high (1,000 p.p.m.) CO_2 concentration and found a similar inversion in the stomatal index of *epf2-1* and *epf2-2* plants (Extended Data Fig. 3). The *epf2* mutant epidermis has been shown to have more non-stomatal cells than WT plants[7,8]. The *epf2* mutants also had more non-stomatal cells at the elevated CO_2 concentration than WT plants (Extended Data Fig. 4a, b). Conversely, plants with a mutation in the related negative-regulatory secreted peptides EPF1 (ref. 9) or EPFL6 (also known as CHALLAH)[21], which also have roles in stomatal development, did not show an inversion of the CO_2-controlled stomatal development response to the elevated CO_2 concentration (Extended Data Fig. 4c, d).

EPF2 belongs to a family of 11 EPF and EPFL peptide proteins, which are predicted to be converted to an active peptide ligand isoform upon cleavage[22–25]. Hence, we tested plants with mutated *SDD1*, which has been shown to be a negative regulator of stomatal development and which encodes an extracellular subtilisin-like serine protease[26]. The stomatal index of the *sdd1-1* mutant was much higher than that of the corresponding C24 WT accession at both the low and elevated CO_2 concentrations (Fig. 3a). The *sdd1-1* mutant showed, on average, a 4% decrease in the stomatal index at the elevated CO_2 concentration compared with the low concentration, similar to the C24 WT background line (Fig. 3a). This result indicates that the protease SDD1 is not, alone, essential for CO_2 control of stomatal development, consistent with studies suggesting that SDD1 does not function in the same pathway as EPF2 (refs 7, 8) and that extracellular proteases that function in the EPF1, EPF2 and STOMAGEN (also known as EPFL9 (refs 23, 24, 27), a positive-regulatory peptide related to EPF1 and EPF2) pathways remain

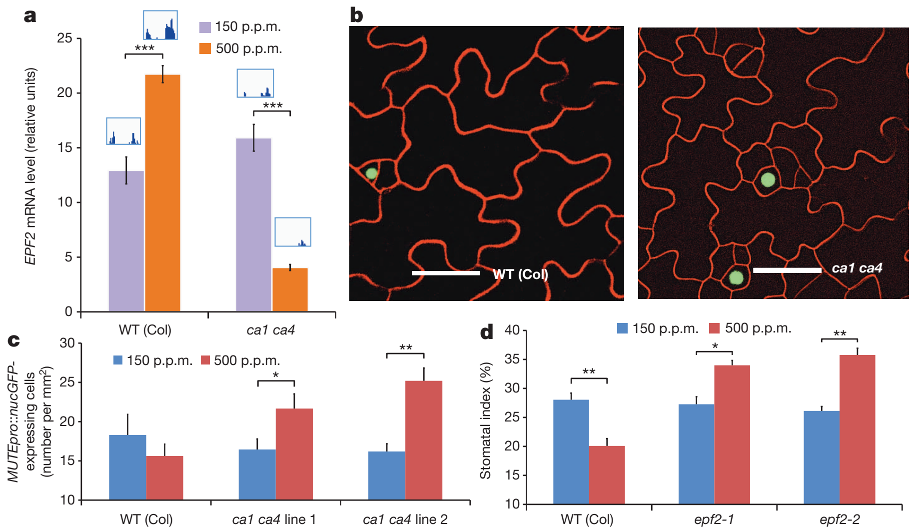


Figure 2 | *EPF2* expression is regulated by CO_2 concentration and is essential for CO_2 control of stomatal development. **a**, *EPF2* messenger RNA levels in developing 5 DAG (days after germination) cotyledons of WT and *ca1 ca4* seedlings, showing induction, at the elevated CO_2 concentration in the WT but not *ca1 ca4*. Levels were normalized to those of the *CLATHRIN* gene. The insets show the normalized RNA-seq expression of *EPF2* exons from an RNA-seq experiment (5 DAG). **b**–**d**, *MUTE* expression correlates with the stomatal development phenotype of the *ca1 ca4* mutant. Confocal images showing *MUTEpro::nucGFP* expression (green) in developing (5 DAG) cotyledons of WT and *ca1 ca4* plants (**b**). Scale bars, 100 μm. Quantitation of *MUTEpro::nucGFP*-expressing cells in the WT and two independent lines in the *ca1 ca4* background, at low and elevated CO_2 concentrations (**c**). **d**, Stomatal index in WT plants and plants carrying either of two independent mutant alleles of *epf2*, at low and elevated CO_2 concentrations, demonstrating that *epf2* mutants show an inversion of the elevated CO_2-mediated control of stomatal development. Error bars, mean ± s.e.m., $n = 10$ in **a** and $n = 20$ in **c** and **d**. ***, $P < 0.00005$; **, $P < 0.005$; *, $P < 0.05$, using ANOVA and Tukey's post-hoc test.

unknown. At present, no environmental signals that clearly mediate the control of stomatal development via the extracellular pro-peptides EPF1, EPF2 and EPFL9 or the protease SDD1 have been identified.

We hypothesized that there is a distinct extracellular protease(s) that mediates CO_2 control of stomatal development. SDD1 belongs to a 56-member subtilisin-like serine protease family (subtilases). Therefore, we pursued proteomic analyses of apoplast proteins in leaves and identified four abundant subtilases (SBT1.7 (also known as ARA12), SBT1.8 (At2g05920), SBT3.13 (At4g21650) and SBT5.2; Extended Data Fig. 5). Because SBT1.7 has been shown to be required for seed mucilage release[28] and SBT3.13 was detected in two of five experiments, we focused on SBT5.2 rather than SBT3.13, SBT1.7 or its closest homologue, SBT1.8. Interestingly, qPCR data from developing cotyledons showed an increase in the abundance of *SBT5.2* transcripts in WT plants after both long term (5 days; Fig. 3b) and short term (4 h; Extended Data Fig. 5f) exposure to the elevated CO_2 concentration. By contrast, the *ca1 ca4* plants failed to show this increase in *SBT5.2* transcript abundance at the elevated CO_2 concentration (Fig. 3b). We named SBT5.2 as CRSP (CO_2 RESPONSE SECRETED PROTEASE). *CRSP* is widely expressed in guard cells and meristemoid- and pavement-cell-enriched samples, as well as in other plant tissues, including high expression in roots[17,29]. Our experiments with a CRSP–VENUS construct showed that CRSP is targeted to the cell wall (Extended Data Fig. 5c, d). We tested the effect on CO_2 control of stomatal development of two T-DNA insertion alleles encoding mutated forms of this extracellular protease (Fig. 3c and Extended Data Figs 1b, 3, 4 and 5e). Interestingly, the two distinct *crsp* mutant alleles (Extended Data Fig. 5e) conferred, on average, deregulation of stomatal development, with more stomata at the elevated CO_2 concentration than at the low concentration (Fig. 3c and Extended Data Figs 1b and 3). Furthermore, when epidermal cell types were analysed individually, the *crsp-1* mutant had more stomata and non-stomatal cells than the WT, which is a similar phenotype to (but not as severe as) the *epf2* mutant (Extended Data Fig. 4a, b), implicating the functions of additional proteases. It should be noted that, similar to *ERECTA*, the wide expression pattern of *CRSP* indicates that the CRSP protein could have additional roles in plant growth and development.

To determine whether the EPF2 pro-peptide can be cleaved by CRSP, we constructed two synthetic peptides spanning the predicted EPF2 cleavage site. We subjected these peptides to *in vitro* proteolytic analyses using *in vitro*-synthesized CRSP protein. CRSP showed robust cleavage of both synthetic EPF2 (synEPF2) peptides *in vitro*, and this cleavage was greatly reduced by the inclusion of protease inhibitors or the mutant form of the CRSP protein (CRSP-1) in the reaction (Extended Data Fig. 6a, e). To test the specificity of CRSP-mediated cleavage, we generated an EPF2 mutant peptide sequence with 7 residue substitutions to mimic a 12-residue sequence that surrounds the cleavage site in STOMAGEN; this mutant was not cleaved by CRSP (Extended Data Fig. 6d). We also tested the synthetic EPF1 and STOMAGEN peptides, and both of these control peptides showed negligible cleavage *in vitro* in the presence of either CRSP or the mutant CRSP-1 (Extended Data Fig. 6b, c). These data support the function of CRSP in the modulation of EPF2 activity.

Several proteomic approaches were unsuccessful at detecting low-abundance EPF1 and EPF2 peptides in cell-wall extracts (see Methods). To further analyse whether EPF2 and CRSP function in the same pathway, we conducted epistasis analyses by generating *crsp epf2* double mutant lines. Double mutant plants did not show clearly additive mutant phenotypes (Extended Data Fig. 7f). We then overexpressed *EPF2* in the WT and *crsp* mutant backgrounds using an oestradiol-inducible system. Analysis of 36 independent lines showed that equivalent quantified levels of *EPF2* overexpression repressed stomatal development to a lesser degree in the *crsp* background than in the WT (Fig. 3d and Extended Data Fig. 7a–e). The partial repression of stomatal density in high-*EPF2*-expressing *crsp* lines, the epistasis analysis and the non-stomatal cell densities implicate the function of additional proteases in EPF2 activation (Extended Data Figs 3, 8 and 9). These data also do not exclude a possible role for CRSP in other stomatal responses. Controls using inducible *EPF1* overexpression showed similar effects on stomatal development in the WT and *crsp* backgrounds (Extended Data Fig. 8).

We have uncovered key elements in a long-sought pathway by which elevated CO_2 concentrations control cell fate and the stomatal development machinery[4]. The results of our study identify new players in CO_2

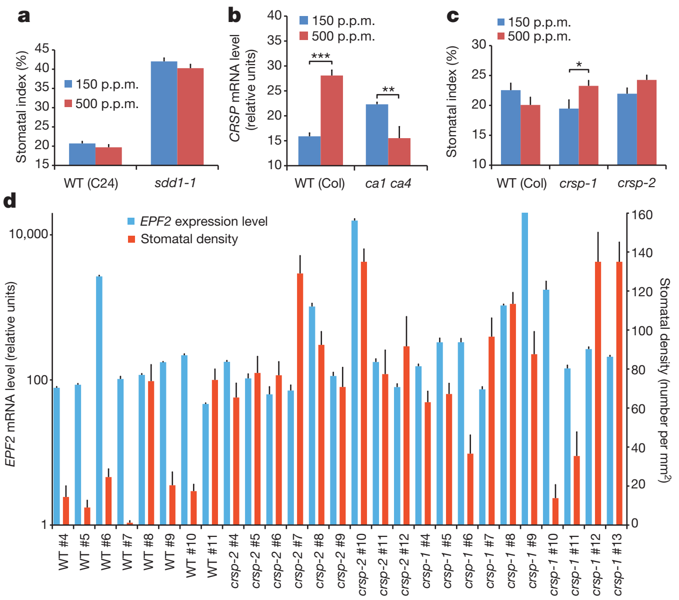


Figure 3 | A CO_2-regulated, secreted subtilisin-like serine protease, CRSP, is a mediator of elevated CO_2 repression of stomatal development. **a**, Stomatal index of the WT (C24) and the *sdd1-1* mutant grown at the low and elevated CO_2 concentrations. **b**, CO_2 control of *CRSP* (*SBT5.2*) mRNA levels in developing (5 DAG) cotyledons of WT (Col) and *ca1 ca4* seedlings grown at low and elevated CO_2 concentrations (qPCR data, with cDNA levels normalized to *CLATHRIN* (At4G24550) expression). **c**, Stomatal index of WT cotyledons and those carrying either of two independent *crsp* alleles at low and elevated CO_2 concentrations. **d**, Quantitation of the effects of *EPF2* transcript levels on the stomatal density of 5 DAG cotyledons in 27 independent lines harbouring the β-oestradiol-inducible *EPF2* overexpression construct in the WT (Col), *crsp-1* and *crsp-2* mutant backgrounds (normalized to *ACTIN 2* expression). For each line, 20 images from 10 cotyledons (2 images per cotyledon; 10 separate seedlings used) were analysed, and RNA was extracted from 10 separate seedlings (see Methods and Extended Data Fig. 7e). Error bars, mean ± s.e.m., $n = 20$ in **a**, **c** and **d** and $n = 10$ in **b**. **b**, **c**, ***, $P < 0.00005$; **, $P < 0.005$; *, $P < 0.05$, using ANOVA and Tukey's post-hoc test.

control of stomatal development: CA1, CA4, CRSP and EPF2. Together, the present findings point to the extracellular protease CRSP, identified here as functioning in the CO_2-controlled stomatal development response, and further suggest that the activity of the negative regulator EPF2 is modulated by CRSP. EPF2 peptides are predicted to be activated by cleavage, thus signalling the repression of stomatal development[7,8,22]. CRSP can cleave EPF2 (Extended Data Fig. 6a, e), and our data provide evidence that CRSP functions in EPF2 signalling to mediate the repression of stomatal development (Fig. 3d and Extended Data Figs 6–8). An inverted CO_2-dependent stomatal development response in *erecta* plants potentially correlates with the preferential binding of EPF2 to the receptor kinase ERECTA[22] (Extended Data Fig. 9).

The finding that the stomatal index is similar in *ca1 ca4* and WT plants at a low CO_2 concentration indicates that additional regulatory mechanisms exist and that CO_2 control is not entirely disrupted in *ca1 ca4* plants. In the absence of the elevated CO_2-mediated modulation of *CRSP* and *EPF2*, competing extracellular signals that promote stomatal development (for example, the STOMAGEN peptide[23,24,27]) might contribute to the inverted CO_2 control of stomatal development found here in the *ca1 ca4*, *epf2* and *crsp* mutants (Figs 1–3). The mechanisms reported here may also aid in understanding the natural variation in stomatal developmental responses to elevated CO_2 concentrations that has been observed in *A. thaliana* and other plant species[30]. Globally, as plants grow and respond to the continuing rise in atmospheric CO_2 concentrations, an understanding of the key genetic players that mediate the CO_2-controlled plant developmental response could become critical for agriculturally relevant efforts aimed at improving water use efficiency or plant heat resistance.

METHODS SUMMARY

Wild type (Col and C24 accessions) and individual mutant seedlings were grown in plant growth chambers (Percival) under identical conditions of light (16 h light:8 h dark cycles; 100 μmol m^{-1} s^{-1}), humidity (80–90%) and temperature (21 °C), with only the CO_2 concentration being varied (low = 150 p.p.m. and elevated = 500 p.p.m. (or 1,000 p.p.m. where noted)). In previous transformant analyses of *ca1 ca4*, YFP fusions of carbonic anhydrases were not used[6], whereas here YFP fusions were used to ascertain developmental-stage-dependent and guard cell expression of carbonic anhydrases. For *MUTE* expression studies, a *MUTEpro::nucGFP*[19] construct was used. It should be noted that absolute stomatal indices and the degree of change in indices varied slightly from experiment to experiment, similar to the findings of previous studies[5], requiring parallel controls and blinded experiments.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 June 2012; accepted 6 May 2014.
Published online 6 July 2014.

1. Bergmann, D. C. & Sack, F. D. Stomatal development. *Annu. Rev. Plant Biol.* **58,** 163–181 (2007).
2. Pillitteri, L. J. & Torii, K. U. Mechanisms of stomatal development. *Annu. Rev. Plant Biol.* **63,** 591–614 (2012).
3. Nadeau, J. A. & Sack, F. D. Control of stomatal distribution on the *Arabidopsis* leaf surface. *Science* **296,** 1697–1700 (2002).
4. Woodward, F. I. Stomatal numbers are sensitive to increases in CO_2 from pre-industrial levels. *Nature* **327,** 617–618 (1987).
5. Gray, J. E. *et al.* The HIC signalling pathway links CO_2 perception to stomatal development. *Nature* **408,** 713–716 (2000).
6. Hu, H. *et al.* Carbonic anhydrases are upstream regulators of CO_2-controlled stomatal movements in guard cells. *Nature Cell Biol.* **12,** 87–93 (2010).
7. Hara, K. *et al.* Epidermal cell density is autoregulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR 2 in *Arabidopsis* leaves. *Plant Cell Physiol.* **50,** 1019–1031 (2009).
8. Hunt, L. & Gray, J. E. The signaling peptide EPF2 controls asymmetric cell divisions during stomatal development. *Curr. Biol.* **19,** 864–869 (2009).
9. Hara, K., Kajita, R., Torii, K. U., Bergmann, D. C. & Kakimoto, T. The secretory peptide gene *EPF1* enforces the stomatal one-cell-spacing rule. *Genes Dev.* **21,** 1720–1725 (2007).
10. Kim, T. H., Bohmer, M., Hu, H., Nishimura, N. & Schroeder, J. I. Guard cell signal transduction network: advances in understanding abscisic acid, CO_2, and Ca^{2+} signaling. *Annu. Rev. Plant Biol.* **61,** 561–591 (2010).
11. Woodward, F. I. Potential impacts of global elevated CO_2 concentrations on plants. *Curr. Opin. Plant Biol.* **5,** 207–211 (2002).
12. Hetherington, A. M. & Woodward, F. I. The role of stomata in sensing and driving environmental change. *Nature* **424,** 901–908 (2003).
13. Masle, J., Gilmore, S. R. & Farquhar, G. D. The *ERECTA* gene regulates plant transpiration efficiency in *Arabidopsis. Nature* **436,** 866–870 (2005).
14. Sellers, P. J. Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science* **275,** 502–509 (1997).
15. Battisti, D. S. & Naylor, R. L. Historical warnings of future food insecurity with unprecedented seasonal heat. *Science* **323,** 240–244 (2009).
16. Yang, Y., Costa, A., Leonhardt, N., Siegel, R. S. & Schroeder, J. I. Isolation of a strong *Arabidopsis* guard cell promoter and its potential role as a research tool. *Plant Methods* **4,** 1 (2008).
17. Pillitteri, L. J., Peterson, K. M., Horst, R. J. & Torii, K. U. Molecular profiling of stomatal meristemoids reveals new component of asymmetric cell division and commonalities among stem cell populations in *Arabidopsis. Plant Cell* **23,** 3260–3275 (2011).
18. Xue, S. *et al.* Central functions of bicarbonate in S-type anion channel activation and OST1 protein kinase in CO_2 signal transduction in guard cell. *EMBO J.* **30,** 1645–1658 (2011).
19. MacAlister, C. A., Ohashi-Ito, K. & Bergmann, D. C. Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. *Nature* **445,** 537–540 (2007).
20. Pillitteri, L. J., Sloan, D. B., Bogenschutz, N. L. & Torii, K. U. Termination of asymmetric cell division and differentiation of stomata. *Nature* **445,** 501–505 (2007).
21. Abrash, E. B. & Bergmann, D. C. Regional specification of stomatal production by the putative ligand CHALLAH. *Development* **137,** 447–455 (2010).
22. Lee, J. S. *et al.* Direct interaction of ligand–receptor pairs specifying stomatal patterning. *Genes Dev.* **26,** 126–136 (2012).
23. Sugano, S. S. *et al.* Stomagen positively regulates stomatal density in *Arabidopsis. Nature* **463,** 241–244 (2010).
24. Kondo, T. *et al.* Stomatal density is controlled by a mesophyll-derived signaling molecule. *Plant Cell Physiol.* **51,** 1–8 (2010).

25. Uchida, N. & Tasaka, M. Regulation of plant vascular stem cells by endodermis-derived EPFL-family peptide hormones and phloem-expressed ERECTA-family receptor kinases. *J. Exp. Bot.* **64,** 5335–5343 (2013).
26. Berger, D. & Altmann, T. A subtilisin-like serine protease involved in the regulation of stomatal density and distribution in *Arabidopsis thaliana*. *Genes Dev.* **14,** 1119–1131 (2000).
27. Ohki, S., Takeuchi, M. & Mori, M. The NMR structure of stomagen reveals the basis of stomatal density regulation by plant peptide hormones. *Nature Commun.* **2,** 512 (2011).
28. Rautengarten, C. *et al.* A subtilisin-like serine protease essential for mucilage release from *Arabidopsis* seed coats. *Plant J.* **54,** 466–480 (2008).
29. Schmid, M. *et al.* A gene expression map of *Arabidopsis thaliana* development. *Nature Genet.* **37,** 501–506 (2005).
30. Woodward, F. I., Lake, J. A. & Quick, W. P. Stomatal development and CO_2: ecological consequences. *New Phytol.* **153,** 477–484 (2002).

Acknowledgements We thank K. Knepper for conducting independent CO_2-dependent stomatal development analyses. We thank A. Ries for help with generating the CA–YFP-fusion complementation lines. We thank D. Bergmann for providing the *epfl6* mutant line and DNA constructs for *MUTEpro*::*nucGFP* expression; K. Torii for providing DNA constructs for *MUTEpro*::*MUTE-GFP* expression, *erecta* mutants and the oestradiol-inducible *EPF* constructs; T. Altmann for providing the *sdd1-1* mutant; and M. Estelle, Y. Zhao, A. Stephan and M. Facette for comments on the manuscript. This project was funded by grants from the National Science Foundation (MCB0918220 and MCB1414339 to J.I.S. and IOS-1025837 to S.C.P.) and the National Institutes of Health (GM060396-ES010337 to J.I.S.), a BAYER-UC Discovery grant (J.I.S.) and a seed grant from the UCSD-SDCSB (GM085764) Systems Biology Center (C.B.E.). A grant from the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences of the US Department of Energy (DE-FG02-03ER15449) to J.I.S. funded complementation and localization analyses.

Author Contributions J.I.S. and C.B.E. designed experiments. C.B.E. conducted experiments (except for mass spectrometry analyses). C.B.E. generated GFP reporter expression and *EPF* overexpression plant lines. H.H. generated genetic constructs and transgenic plant lines for *ca1 ca4* mutant YFP-fusion lines. Mass spectrometry analyses were conducted by M.G., J.C.A. and S.C.P. J.I.S. proposed the project. The manuscript was written by C.B.E. and J.I.S.

Author Information The raw data from three independent biological replicates in RNA-seq experiments have been deposited in the BioProject database under accession number PRJNA218542. The mass spectrometry proteomics data have been deposited in the Proteomics Identification Database (PRIDE) under accession numbers PXD000692, PXD000693 and PXD000956. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.I.S. (jischroeder@ucsd.edu).

LETTER

doi:10.1038/nature13557

Fructose-1,6-bisphosphatase opposes renal carcinoma progression

Bo Li[1], Bo Qiu[1], David S. M. Lee[1,2], Zandra E. Walton[1,3], Joshua D. Ochocki[1], Lijoy K. Mathew[1,2], Anthony Mancuso[1,3,4], Terence P. F. Gade[4], Brian Keith[1,3], Itzhak Nissim[5,6] & M. Celeste Simon[1,2,7]

Clear cell renal cell carcinoma (ccRCC), the most common form of kidney cancer[1], is characterized by elevated glycogen levels and fat deposition[2]. These consistent metabolic alterations are associated with normoxic stabilization of hypoxia-inducible factors (HIFs)[3] secondary to von Hippel–Lindau (*VHL*) mutations that occur in over 90% of ccRCC tumours[4]. However, kidney-specific *VHL* deletion in mice fails to elicit ccRCC-specific metabolic phenotypes and tumour formation[5], suggesting that additional mechanisms are essential. Recent large-scale sequencing analyses revealed the loss of several chromatin remodelling enzymes in a subset of ccRCC (these included polybromo-1, SET domain containing 2 and BRCA1-associated protein-1, among others)[6–9], indicating that epigenetic perturbations are probably important contributors to the natural history of this disease. Here we used an integrative approach comprising pan-metabolomic profiling and metabolic gene set analysis and determined that the gluconeogenic enzyme fructose-1,6-bisphosphatase 1 (FBP1)[10] is uniformly depleted in over six hundred ccRCC tumours examined. Notably, the human *FBP1* locus resides on chromosome 9q22, the loss of which is associated with poor prognosis for ccRCC patients[11]. Our data further indicate that FBP1 inhibits ccRCC progression through two distinct mechanisms. First, FBP1 antagonizes glycolytic flux in renal tubular epithelial cells, the presumptive ccRCC cell of origin[12], thereby inhibiting a potential Warburg effect[13,14]. Second, in pVHL (the protein encoded by the *VHL* gene)-deficient ccRCC cells, FBP1 restrains cell proliferation, glycolysis and the pentose phosphate pathway in a catalytic-activity-independent manner, by inhibiting nuclear HIF function via direct interaction with the HIF inhibitory domain. This unique dual function of the FBP1 protein explains its ubiquitous loss in ccRCC, distinguishing FBP1 from previously identified tumour suppressors that are not consistently mutated in all tumours[6,7,15].

We performed pan-metabolomic analysis on 20 primary human ccRCC tumours and matching normal kidney tissues. Levels of metabolites involved in glycolysis, gluconeogenesis and glucose-related sugar metabolism were highly elevated in tumours, suggesting that reprogramming of glucose metabolism is critical for ccRCC progression (Extended Data Fig. 1a). Furthermore, metabolic gene set analysis[16] of The Cancer Genome Atlas ccRCC RNA-sequencing data indicated that the carbohydrate storage group was the most significantly underexpressed gene set in ccRCC tumours (Fig. 1a), including three genes controlling renal gluconeogenesis[17] (glucose-6-phosphatase, catalytic subunit (*G6PC*), phosphoenolpyruvate carboxykinase 1 (*PCK1*), and *FBP1*) (Extended Data Fig. 1b–c). Elevated HIF activity in ccRCC tumours stimulates aerobic glycolysis by increasing the expression of glycolytic genes, including phosphoglycerate kinase 1 (*PGK1*) and lactate dehydrogenase A (*LDHA*), and shunting glycolytic flux away from the TCA cycle by activating pyruvate dehydrogenase kinase 1 (*PDK1*)[3]. However, our integrative analyses identified suppression of gluconeogenesis as an additional component of glucose regulation in ccRCCs. Next, we determined that the rate-limiting gluconeogenic enzyme FBP1 was inhibited at the level of protein accumulation in almost 100% of ccRCC tumours examined ($n > 200$, Fig. 1b and Extended Data Fig. 2a–c) compared to normal kidney tissue. Similar results were observed for hepatocellular carcinomas and normal liver tissue (Extended Data Fig. 2d, e). FBP1 inhibition is not mediated by ccRCC-associated HIF activation, because *HIF1A* (that is, the mRNA that codes for HIF-1α) ablation failed to de-repress FBP1 expression in RCC4 ccRCC cells (Extended Data Fig. 2f). Moreover, HK-2 cells, from proximal tubule epithelial cells (where ccRCCs appear to arise[12]) exhibited HIF-1α-dependent induction of FBP1 under hypoxia (Extended Data Fig. 2g, h). Compared to FBP1, the other two gluconeogenic enzymes G6PC and PCK1 were either modestly suppressed (G6PC), or exhibited no consistent change (PCK1) in ccRCC tumours (Extended Data Fig. 3a, b). Notably, the glycolytic enzyme phosphofructokinase (liver type, PFKL), which functionally antagonizes FBP1 in glycolysis (Extended Data Fig. 1c), was expressed at equal levels in ccRCC and normal kidney tissues (Extended Data Fig. 3c). In addition, lower *FBP1* expression correlates significantly with advanced tumour stage and worse patient prognosis (Fig. 1c, d), whereas *PFKL* expression does not (Extended Data Fig. 3d, e), suggesting that FBP1 may have novel, non-enzymatic function(s).

To investigate functional roles for FBP1 in ccRCC progression, we ectopically expressed FBP1 in 786-O ccRCC tumour cells to levels observed in HK-2 proximal tubule cells (Extended Data Fig. 4a). FBP1 expression significantly inhibited 2D culture (Fig. 1e), anchorage-independent (Extended Data Fig. 4b), and xenograft tumour growth (Fig. 1f and Extended Data Fig. 4c). Similarly, enforced FBP1 expression inhibited growth of RCC10 and 769-P ccRCC cells (Extended Data Fig. 4d, e) and A549 lung cancer cells preferentially under hypoxia (Extended Data Fig. 4f, g). These results demonstrate that FBP1 can suppress ccRCC and other tumour cell growth, an effect significantly pronounced when coupled with HIF activation. In HK-2 cells, FBP1 depletion, but not G6PC ablation or ectopic PFKL expression, was sufficient to promote HK-2 cell growth (Fig. 1g and Extended Data Fig. 4h–j).

Since FBP1 is the rate-limiting enzyme in gluconeogenesis[10], we manipulated FBP1 expression in renal cells and measured glucose metabolism. FBP1 inhibition increased glucose uptake and lactate secretion in HK-2 cells cultured in 10 mM glucose, (Fig. 2a), an effect augmented by lowering glucose levels to 1 mM (Extended Data Fig. 5a, b). To assess glycolytic flux, we performed isotopomer distribution analysis using [1,2-^{13}C]glucose as the tracer, which produces glycolytic and TCA intermediates containing two ^{13}C atoms (M2 species), as well as corresponding M1 species containing one ^{13}C atom from the pentose phosphate pathway (PPP; Extended Data Fig. 5c). We observed elevated M2 enrichment of four TCA intermediates (malate, aspartate, glutamate and citrate) in FBP1-depleted HK-2 cells (Fig. 2b, c). In contrast, G6PC inhibition failed to promote glucose-lactate turnover (data not shown), suggesting that FBP1, but not G6PC, is a critical regulator of glucose metabolism in renal cells. Consistent with this result, ectopic FBP1 expression in a

[1]Abramson Family Cancer Research Institute, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [2]Howard Hughes Medical Institute, Philadelphia, Pennsylvania 19104, USA. [3]Department of Cancer Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [4]Department of Radiology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [5]Department of Pediatrics, Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [6]Division of Child Development and Metabolic Disease, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA. [7]Department of Cell and Developmental Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

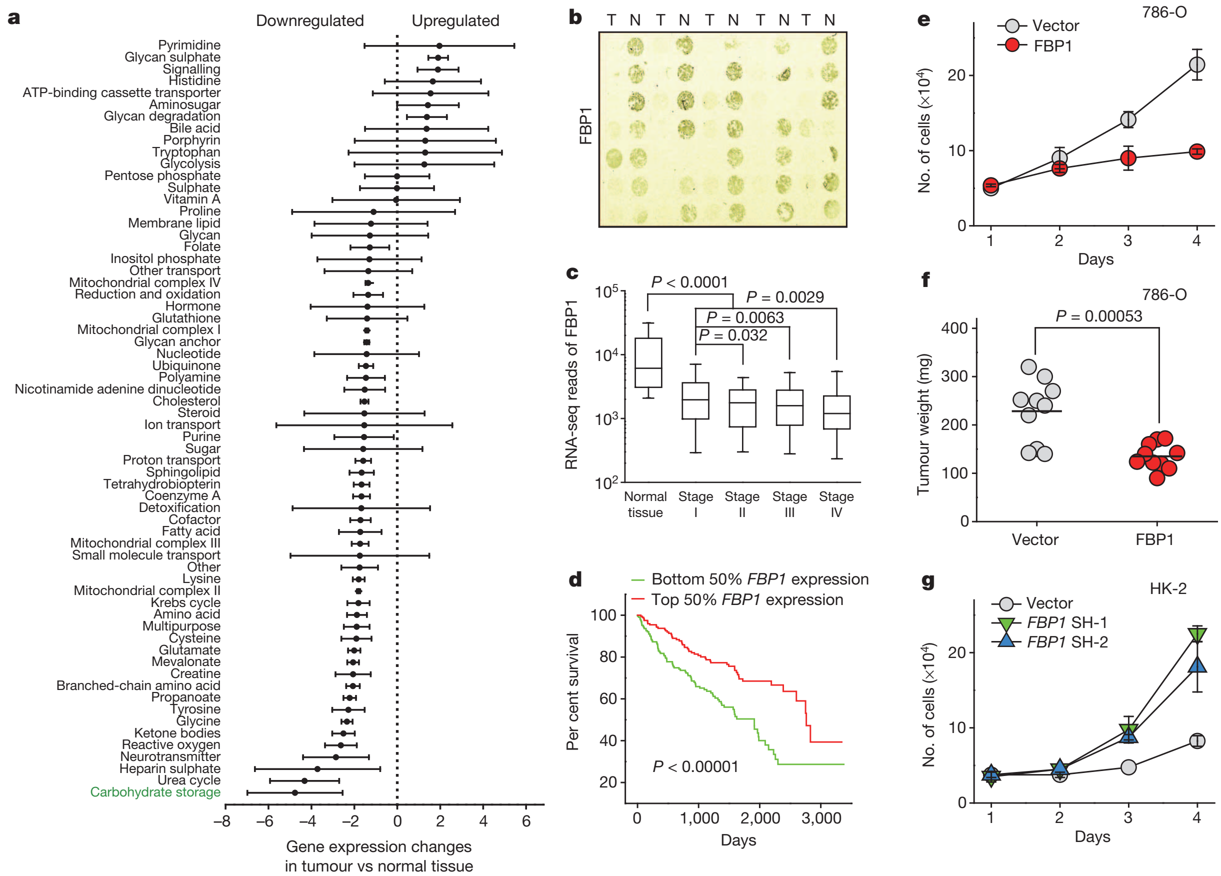


Figure 1 | Integrative analyses reveal that FBP1 is ubiquitously depleted and exhibits tumour-suppressive functions in ccRCC. **a**, Metabolic gene set analysis of RNA-sequencing (RNA-seq) data provided by The Cancer Genome Atlas (TCGA) ccRCC project. A total of 480 ccRCC tumour and 69 adjacent normal tissues were included and 2,752 genes encoding all known human metabolic enzymes and transporters were classified according to the Kyoto Encyclopaedia of Genes and Genomes. Generated metabolic gene sets were ranked based on their median fold expression changes in ccRCC tumour vs normal tissue, and plotted as median ± median absolute deviation. $^*P < 0.01$. **b**, Immunohistochemistry staining of a representative kidney tissue microarray with FBP1 antibody. T, ccRCC tumours; N, adjacent normal kidney. **c**, Normalized RNA-seq reads of *FBP1* in 69 normal kidneys and 480 ccRCC tumours grouped into stage I–IV by TCGA. **d**, Kaplan–Meier survival curve of 429 ccRCC patients enrolled in the TCGA database. Patients were equally divided into two groups (top and bottom 50% *FBP1* expression) based on *FBP1* mRNA levels in their tumours. **e**, Growth of 786-O ccRCC cells in low serum medium (1% FBS), with or without ectopic FBP1 expression. **f**, Xenograft tumour growth of 786-O cells with or without ectopic FBP1 expression (each group includes ten tumours from five mice). End-point tumour weights were measured and plotted. The horizontal bars represent means of tumour weights within each group. **g**, Growth of human HK-2 proximal renal tubule cells with or without FBP1 inhibition in 1% serum medium. SH-1 and SH-2 represent two different *FBP1* short hairpin RNAs. Values represent mean ± standard deviation (s.d.) (four technical replicates, from two independent experiments).

VHL-deficient ccRCC cell line (RCC10) reduced glucose uptake, lactate secretion and glucose-derived TCA cycle intermediates (Fig. 2d, e). Reduced glucose-dependent TCA flux is known to increase anaplerotic glutamine flux[18], and we also observed elevated glutamine uptake and enrichment of glutamine-derived TCA cycle intermediates (M4 species) in [U-^{13}C]glutamine-labelled RCC10 cells expressing FBP1 (Extended Data Fig. 5d–f).

Pan-metabolomic analyses of ccRCC tumours revealed a marked elevation of reduced glutathione (G-SH)[19] (Extended Data Fig. 5g). G-SH synthesis requires the reduced form of nicotinamide adenine dinucleotide phosphate (NADPH), generated primarily though the PPP in human cells[19] (Extended Data Fig. 5c). Consistent with increased PPP flux, ccRCC tumours display significant accumulation of G-SH and PPP-related metabolites (Extended Data Fig. 5g), an effect partially recapitulated in FBP1-depleted HK-2 cells (Extended Data Fig. 5h–j). Conversely, FBP1 re-expression in RCC10 cells significantly reduced NADPH levels and PPP flux (Fig. 2f and Extended Data Fig. 5k, l). Furthermore, FBP1-mediated changes in PPP flux (Extended Data Fig. 5j and 5l) were comparable to the changes in glucose 6-phosphate (G6P; Extended Data Fig. 5m, n), the entry metabolite of the PPP pathway (Extended Data Fig. 5c), suggesting that FBP1 affects PPP flux primarily through regulating glycolysis. Notably, the ability of FBP1 to reduce glycolysis and NADPH levels was completely abolished in RCC10-VHL cells (Fig. 2d, f), that is, cells where wild-type pVHL was reintroduced into RCC10 cells to exclude normoxic HIF expression (Extended Data Fig. 5o), indicating that HIF proteins are required for FBP1-mediated effects on glucose metabolism in ccRCC tumour cells.

To investigate a mechanistic link between FBP1 expression and HIF activity, we employed two *VHL*-deficient ccRCC lines, RCC4 and RCC10, which express both HIF-1α and HIF-2α (Extended Data Fig. 6a). HIF-1α and HIF-2α are induced at different stages of ccRCC and have cooperating and contrasting roles in tumour progression[3]. HIF-1α and HIF-2α function by binding hypoxia response elements (HREs) within target genes, including those modulating cellular metabolism[20]. Notably, ectopic FBP1 expression suppressed HIF activity (Fig. 3a) and promoted oxygen consumption in RCC4 and RCC10 cells (Extended Data Fig. 6b, c). Furthermore, FBP1 expression in RCC10 cells restored pyruvate dehydrogenase activity (Extended Data Fig. 6d–f), which was otherwise inhibited by HIF-1α (ref. 3). Conversely, FBP1 ablation enhanced HIF activity in RCC10 cells, which express detectable levels of FBP1 (Fig. 3a and Extended Data Fig. 6a). This inverse correlation between *FBP1* expression and HIF activity was recapitulated in primary ccRCC tumours (Fig. 3b). In

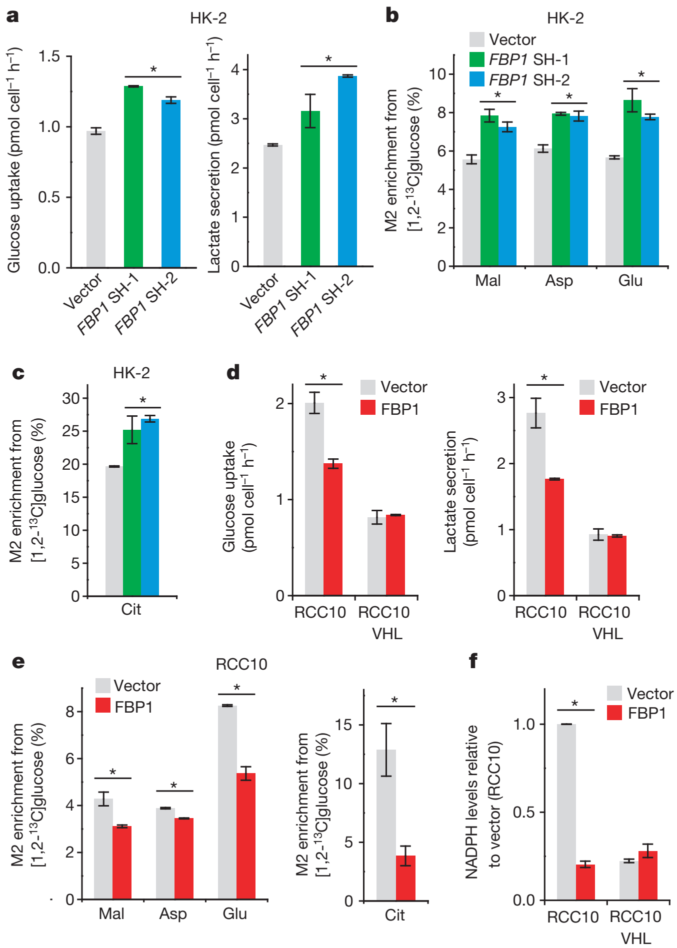


Figure 2 | FBP1 regulates glycolysis and NADPH levels. **a**, Glucose uptake and lactate secretion in HK-2 cells with or without FBP1 depletion. SH-1 and SH-2 represent two different *FBP1* short hairpin RNAs. **b**, **c**, M2 isotopomer distribution of indicated metabolites (**b**) and citrate (**c**) in HK-2 cells with or without FBP1 ablation, labelled with [1,2-^{13}C]glucose. M2 enrichment represents the mole per cent excess of M2 species above natural abundance. **d**, Glucose uptake and lactate secretion in vector control or FBP1-expressing RCC10 and RCC10-VHL cells. RCC10-VHL cells are RCC10 cells where wild-type pVHL has been reintroduced. **e**, M2 isotopomer distribution of indicated metabolites in vector control or FBP1-expressing RCC10 cells, labelled with [1,2-^{13}C]glucose. **f**, Relative NADPH levels in RCC10 and RCC10-VHL cells as indicated in **d**. Values represent mean ± s.d. (three experimental replicates). *$P < 0.05$.

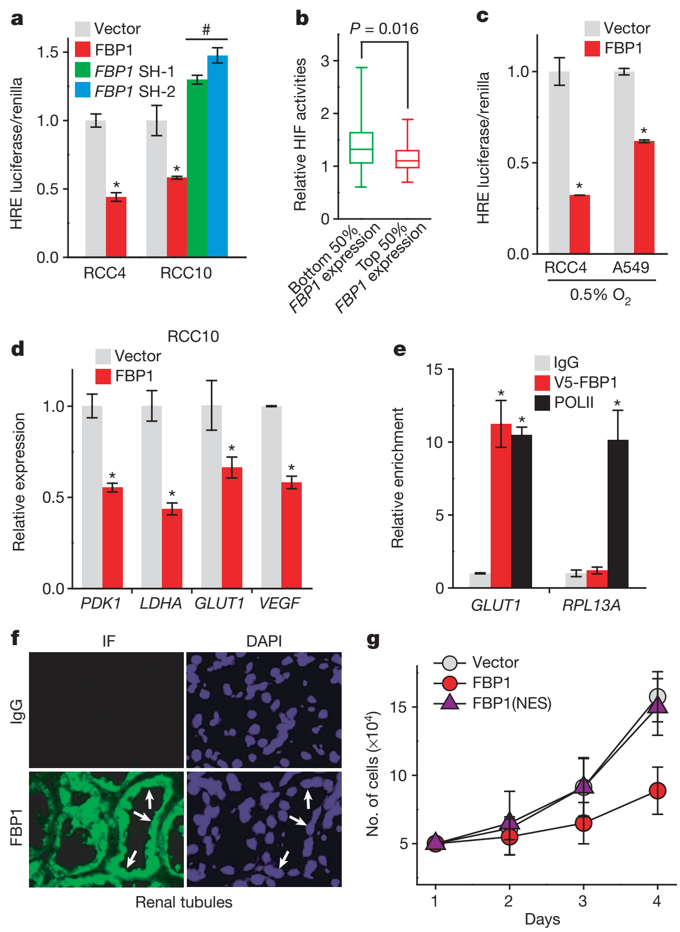


Figure 3 | FBP1 inhibits HIF activity in the nucleus. **a**, HIF reporter activity measured in RCC4 and RCC10 cells transfected with pHRE-luciferase, in the presence of vector, *FBP1* cDNA, or two different *FBP1* short hairpin RNAs (SH-1 or SH-2). Transfection efficiencies were normalized to co-transfected pRenilla-luciferase. The hash symbol represents a significant increase over vector control. **b**, A total of 480 ccRCC tumours from TCGA database were equally divided into two groups (top and bottom 50% *FBP1* expression) based on *FBP1* mRNA levels, and their relative HIF activities were quantified and plotted as described in the Methods. **c**, HIF reporter activity in hypoxic RCC4 and A549 cells (0.5% O_2) with or without ectopic FBP1 expression. **d**, qRT–PCR analysis of HIF target genes in vector control or FBP1-expressing RCC10 cells. **e**, ChIP assays evaluating the chromatin binding of FBP1 to HREs in the *GLUT1* promoter, or to a non-hypoxia responsive region of the *RPL13A* locus. RNA polymerase II (POLII) antibody was used as a positive control. IgG, isotype-matched immunoglobulin G; V5-FBP1, V5-tagged FBP1. **f**, Immunofluorescent staining of primary human kidney tissue (tubular region) with FBP1 antibody. Arrows point to three representative sites with nuclear FBP1. Rabbit IgG was used as a negative control, and DAPI is a fluorescent nuclear dye. **g**, Growth of vector control, FBP1 or FBP1(NES)-expressing RCC10 cells cultured in 1% serum. FBP1(NES) refers to FBP1 linked to a C-terminal nuclear export sequence. Error bars represent s.d. (three experimental replicates) except in **e**, which indicates standard error of the mean (three technical replicates from a representative experiment). *$P < 0.05$; #$P < 0.05$.

contrast, *G6PC* expression did not correlate with HIF activity in ccRCC cells or tumour tissues (Extended Data Fig. 6g, h). FBP1 also inhibited HIF activity in A549 lung cancer cells cultured at 0.5% O_2 (Fig. 3c), demonstrating that this effect is not specific to renal cells. Moreover, FBP1 expression reduced canonical HIF target (*PDK1*, *LDHA*, glucose transporter 1 (*GLUT1*, also known as *SLC2A1*) and vascular endothelial growth factor (*VEGF*)) mRNA levels in RCC4, RCC10 and hypoxic A549 cells, but not in normoxic RCC10-VHL cells (Fig. 3d and Extended Data Fig. 6i–k). Chromatin immunoprecipitation (ChIP) analyses indicated that FBP1 was enriched at the HREs of *PDK1*, *LDHA*, *GLUT1* and *VEGF* promoters, but not at the promoter of ribosomal protein L13a (*RPL13A*), which is non-responsive to hypoxia (Fig. 3e and Extended Data Fig. 7a). ChIP-reChIP analyses revealed co-localization of HIF-1α and FBP1 at these HREs (Extended Data Fig. 7b), suggesting that FBP1 directly inhibits HIF in the nucleus, a conclusion supported by cellular fractionation and immunofluorescent staining of primary human kidney tissue (Fig. 3f and Extended Data Fig. 7c, d). Furthermore, a nucleus-excluded form of FBP1 (FBP1(NES)) containing a potent nuclear export sequence[21] fused to the FBP1 carboxy terminus (Extended Data Fig. 7e) failed to inhibit HIF target gene expression as efficiently as wild-type FBP1 (Extended Data Fig. 7f). Expression of FBP1(NES) neither suppressed RCC10 cell growth nor altered glucose-lactate turnover, in contrast to wild-type FBP1 (Fig. 3g and Extended Data Fig. 7g, h). Collectively, these data demonstrate that nuclear FBP1 is required for inhibiting HIF and glucose metabolism in *VHL*-deficient ccRCC cells (Extended Data Fig. 7i).

To determine whether FBP1 enzymatic activity is required to inhibit HIF, we expressed a previously described, catalytically inactive FBP1(G260R)

mutant[22,23] (Fig. 4a and Extended Data Fig. 8a, b) in RCC10 and 786-O cells (Extended Data Fig. 8a, c). FBP1(G260R) inhibited cell growth to a comparable level of wild-type FBP1 in 10 mM glucose (Fig. 4b and Extended Data Fig. 8d). FBP1(G260R) also inhibited glucose metabolism, NADPH production and HIF target gene expression to the same extent as wild-type FBP1 in RCC10 cells (Extended Data Fig. 8e–h). These results suggest that FBP1 interferes with HIF function through a catalytic-activity-independent mechanism. In normoxic RCC10-VHL cells (low HIF activity), the ability of the FBP1(G260R) mutant to inhibit cell growth, glucose metabolism, NADPH production and HIF target gene expression was abolished (Extended Data Fig. 8c, i–m), further confirming that FBP1 impacts ccRCC cell metabolism and growth by regulating HIF, independent of its enzymatic activity. Nevertheless, wild-type FBP1 suppressed RCC10 and 786-O (Extended Data Fig. 9a, b) cell growth more potently than the G260R mutant under low-glucose conditions, presumably because the enzymatic inhibition of glycolysis by wild-type FBP1 is more profound when glucose supply becomes limited.

To explore the molecular mechanism(s) whereby FBP1 inhibits HIF activity, we separated the FBP1 protein into an N-terminal regulatory (R) domain containing allosteric regulatory sites and a C-terminal (C) domain containing the catalytic site (Fig. 4a). Notably, ectopically expressing the FBP1 R domain in RCC10, RCC4 and 786-O cells was sufficient to inhibit HIF activity, whereas expressing the C domain was not (Fig. 4c and Extended Data Fig. 9c, d). As 786-O cells express HIF-2α but not functional HIF-1α (ref. 24), we conclude that FBP1 inhibits both HIF-1α and HIF-2α, presumably through a similar mechanism. To further map critical FBP1 regions for HIF recognition, we systematically deleted each exon from full-length FBP1 (Extended Data Fig. 9e). All seven FBP1 truncations exhibited minimal catalytic activity (Extended Data Fig. 9f), whereas only the N-terminal exon 1 and exon 2 truncations significantly lost their ability to inhibit HIF (Extended Data Fig. 9g).

We further demonstrated the association of FBP1 and HIF-1α by co-immunoprecipitating epitope-tagged and/or endogenous proteins from 293T or RCC10 cell lysates (Fig. 4d and Extended Data Fig. 9h, i). FBP1 also associated with HIF-2α (Fig. 4e and Extended Data Fig. 10a, b), but not with PHD2 or FIH1, two well-documented HIF-α regulators[3] (Extended Data Fig. 10c). Interestingly, GST pull-down assays revealed that HIF-1α or HIF-2α proteins bound directly to full-length FBP1 (Fig. 4f and Extended Data Fig. 10d), and the interaction between HIF-1α and FBP1 is dependent on FBP1 exon 1 or exon 2 (Fig. 4g). Furthermore, FBP1 associates with the relatively uncharacterized HIF inhibitory domain[25] (Fig. 4h and Extended Data Fig. 10e). To examine whether FBP1 inhibits HIF activity through inhibitory domain recognition, we replaced the HIF-α DNA binding domain with a GAL4 DNA binding domain (GBD) and performed GAL4 transactivation assays (Extended Data Fig. 10e). Consistent with HIF reporter assays (Figs 3a, c, 4c and Extended Data Fig. 9g), FBP1 suppressed full-length HIF-1α(GBD) activity by approximately 50% (Extended Data Fig. 10f, red column). Importantly, removal of the HIF-1α inhibitory domain largely relieved the FBP1 inhibitory effect (Extended Data Fig. 10f). In HIF-2α, the critical region mediating FBP1 inhibition extended to the entire C terminus (Extended Data Fig. 10g). Therefore, FBP1 suppresses HIF-1α and HIF-2α activity by interacting with their C-terminal regions, especially the inhibitory domain motif.

Apart from *VHL* loss, ccRCCs exhibit remarkable genetic heterogeneity[26]. Recent large-scale analyses identified frequent mutations in three epigenetic genes — *PBRM1*, *SETD2* and *BAP1* — all of which reside in a 43-megabase region on chromosome 3p that encompasses *VHL* (refs 6–9). Histologically, ccRCC is characterized by the clear-cell phenotype resulting from glycogen and lipid accumulation[2], suggesting that metabolic perturbations are a defining feature of these tumours. Here we demonstrate that the gluconeogenic enzyme FBP1 is ubiquitously depleted in ccRCCs, consistent with our previous copy number analyses[27]. Moreover, FBP1 exhibits dual tumour-suppressive functions mediated by two separate domains, explaining the universal loss of FBP1 expression in ccRCC tumours. Collectively, our data reveal an intriguing regulatory relationship between FBP1 and hypoxic responses in renal carcinoma, which has implications for the metabolic regulation of all gluconeogenic tissues (Supplementary Discussion).

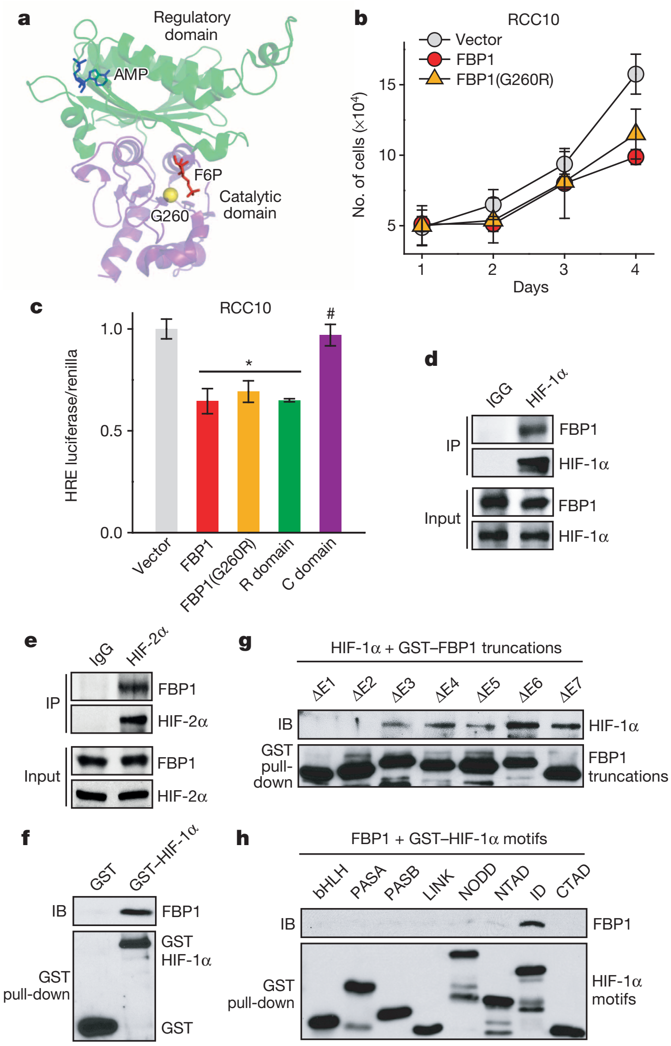


Figure 4 | FBP1 inhibits HIF independent of its enzymatic activity, through direct interaction with a HIF-α inhibitory domain. **a**, Crystal structure (Protein Data Bank accession 1EYJ) of porcine FBP1 in complex with adenosine monophosphate (AMP, blue) and fructose 6-phosphate (F6P, red). The N-terminal regulatory domain of FBP1 is coloured in green, and the C-terminal catalytic domain is coloured in violet. The G260 residue is highlighted in yellow. **b**, Growth of vector control, FBP1 or FBP1(G260R)-expressing RCC10 cells in 1% serum medium. **c**, HIF reporter activity in vector control RCC10 cells, or RCC10 cells expressing FBP1, FBP1(G260R), the regulatory domain of FBP1 (R domain), and catalytic domain of FBP1 (C domain). The hash symbol represents a significant de-repression of HIF reporter activity relative to wild-type FBP1. RCC10 cell lysates were immunoprecipitated with IgG, HIF-1α antibody (**d**), or HIF-2α antibody (**e**) and blotted for endogenous FBP1. IP, immunoprecipitate. **f**, GST pull-down analysis between recombinant FBP1 and recombinant GST or GST-tagged HIF-1α. IB, immunoblot. **g**, GST pull-down analysis between recombinant HIF-1α and recombinant GST-tagged, FBP1 exon truncations. **h**, GST pull-down analysis between recombinant FBP1 and GST-tagged HIF-1α motifs. Values represent mean ± s.d. (three experimental replicates). $*P < 0.01$; $\#P < 0.01$.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 December 2013; accepted 6 June 2014.

Published online 20 July; corrected online 10 September 2014 (see full-text HTML version for details).

1. Rini, B. I., Campbell, S. C. & Escudier, B. Renal cell carcinoma. *Lancet* **373,** 1119–1132 (2009).

2. Valera, V. A. & Merino, M. J. Misdiagnosis of clear cell renal cell carcinoma. *Nature Rev. Urol.* **8,** 321–333 (2011).
3. Keith, B., Johnson, R. S. & Simon, M. C. HIF1α and HIF2α: sibling rivalry in hypoxic tumour growth and progression. *Nature Rev. Cancer* **12,** 9–22 (2012).
4. Nickerson, M. L. *et al.* Improved identification of von Hippel–Lindau gene alterations in clear cell renal tumors. *Clin. Cancer Res.* **14,** 4726–4734 (2008).
5. Rankin, E. B., Tomaszewski, J. E. & Haase, V. H. Renal cyst development in mice with conditional inactivation of the von Hippel-Lindau tumor suppressor. *Cancer Res.* **66,** 2576–2583 (2006).
6. Sato, Y. *et al.* Integrated molecular analysis of clear-cell renal cell carcinoma. *Nature Genet.* **45,** 860–867 (2013).
7. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **499,** 43–49 (2013).
8. Dalgliesh, G. L. *et al.* Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463,** 360–363 (2010).
9. Varela, I. *et al.* Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature* **469,** 539–542 (2011).
10. Tejwani, G. A. Regulation of fructose-bisphosphatase activity. *Adv. Enzymol.* **54,** 121–194 (1983).
11. Moore, L. E. *et al.* Genomic copy number alterations in clear cell renal carcinoma: associations with case characteristics and mechanisms of VHL gene inactivation. *Oncogenesis* **1,** e14 (2012).
12. Cohen, H. T. & McGovern, F. J. Renal-cell carcinoma. *N. Engl. J. Med.* **353,** 2477–2490 (2005).
13. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324,** 1029–1033 (2009).
14. DeBerardinis, R. J. & Thompson, C. B. Cellular metabolism and disease: what do metabolic outliers teach us? *Cell* **148,** 1132–1144 (2012).
15. Hakimi, A. A. *et al.* Adverse outcomes in clear cell renal cell carcinoma with mutations of 3p21 epigenetic regulators BAP1 and SETD2: a report by MSKCC and the KIRC TCGA research network. *Clin. Cancer Res.* **19,** 3259–3267 (2013).
16. Possemato, R. *et al.* Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476,** 346–350 (2011).
17. Gerich, J. E., Meyer, C., Woerle, H. J. & Stumvoll, M. Renal gluconeogenesis: its importance in human glucose homeostasis. *Diabetes Care* **24,** 382–391 (2001).
18. Metallo, C. M. *et al.* Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. *Nature* **481,** 380–384 (2012).
19. Salway, J. G. *Metabolism at a Glance* 3rd edn (Blackwell, 2004).
20. Majmundar, A. J., Wong, W. J. & Simon, M. C. Hypoxia-inducible factors and the response to hypoxic stress. *Mol. Cell* **40,** 294–309 (2010).
21. Wen, W., Meinkoth, J. L., Tsien, R. Y. & Taylor, S. S. Identification of a signal for rapid export of proteins from the nucleus. *Cell* **82,** 463–473 (1995).
22. Åsberg, C. *et al.* Fructose 1,6-bisphosphatase deficiency: enzyme and mutation analysis performed on calcitriol-stimulated monocytes with a note on long-term prognosis. *J. Inherit. Metab. Dis.* **33** (suppl. 3), 113–121 (2010).
23. Choe, J. Y., Fromm, H. J. & Honzatko, R. B. Crystal structures of fructose 1,6-bisphosphatase: mechanism of catalysis and allosteric inhibition revealed in product complexes. *Biochemistry* **39,** 8565–8574 (2000).
24. Shen, C. *et al.* Genetic and functional studies implicate HIF1α as a 14q kidney cancer suppressor gene. *Cancer Discov.* **1,** 222–235 (2011).
25. Jiang, B. H., Zheng, J. Z., Leung, S. W., Roe, R. & Semenza, G. L. Transactivation and inhibitory domains of hypoxia-inducible factor 1alpha. Modulation of transcriptional activity by oxygen tension. *J. Biol. Chem.* **272,** 19253–19260 (1997).
26. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366,** 883–892 (2012).
27. Dondeti, V. R. *et al.* Integrative genomic analyses of sporadic clear cell renal cell carcinoma define disease subtypes and potential new therapeutic targets. *Cancer Res.* **72,** 112–121 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank Y. Daikhin, O. Horyn and Ilana Nissim for assistance with the isotopomer enrichment analysis in the Metabolomic Core facility, Children's Hospital of Philadelphia. We also thank J. Tobias for help with processing the TCGA RNA-sequencing data. This work was supported by the Howard Hughes Medical Institute, NIH Grant CA104838 to M.C.S. and DK053761 to I.N. M.C.S. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions B.L., I.N. and M.C.S. designed this study. B.L., B.Q., D.S.M.L., Z.E.W. and J.D.O. performed the experiments. B.L., L.K.M., A.M., T.P.F.G., I.N. and M.C.S. analysed data. B.L., I.N., B.K. and M.C.S. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.C.S. (celeste2@mail.med.upenn.edu).

doi:10.1038/nature13466

Metastasis-suppressor transcript destabilization through TARBP2 binding of mRNA hairpins

Hani Goodarzi[1], Steven Zhang[1], Colin G. Buss[1], Lisa Fish[1], Saeed Tavazoie[2] & Sohail F. Tavazoie[1]

Aberrant regulation of RNA stability has an important role in many disease states[1,2]. Deregulated post-transcriptional modulation, such as that governed by microRNAs targeting linear sequence elements in messenger RNAs, has been implicated in the progression of many cancer types[3–7]. A defining feature of RNA is its ability to fold into structures. However, the roles of structural mRNA elements in cancer progression remain unexplored. Here we performed an unbiased search for post-transcriptional modulators of mRNA stability in breast cancer by conducting whole-genome transcript stability measurements in poorly and highly metastatic isogenic human breast cancer lines. Using a computational framework that searches RNA sequence and structure space[8], we discovered a family of GC-rich structural *cis*-regulatory RNA elements, termed sRSEs for structural RNA stability elements, which are significantly overrepresented in transcripts displaying reduced stability in highly metastatic cells. By integrating computational and biochemical approaches, we identified TARBP2, a double-stranded RNA-binding protein implicated in microRNA processing, as the *trans* factor that binds the sRSE family and similar structural elements—collectively termed TARBP2-binding structural elements (TBSEs)—in transcripts. TARBP2 is overexpressed in metastatic cells and metastatic human breast tumours and destabilizes transcripts containing TBSEs. Endogenous TARBP2 promotes metastatic cell invasion and colonization by destabilizing amyloid precursor protein (*APP*) and *ZNF395* transcripts, two genes previously associated with Alzheimer's and Huntington's disease, respectively. We reveal these genes to be novel metastasis suppressor genes in breast cancer. The cleavage product of APP, extracellular amyloid-α peptide, directly suppresses invasion while ZNF395 transcriptionally represses a pro-metastatic gene expression program. The expression levels of *TARBP2*, *APP* and *ZNF395* in human breast carcinomas support their experimentally uncovered roles in metastasis. Our findings establish a non-canonical and direct role for TARBP2 in mammalian gene expression regulation and reveal that regulated RNA destabilization through protein-mediated binding of mRNA structural elements can govern cancer progression.

Gene expression studies, in principle, measure steady-state transcript levels. However, such measurements obscure the role of dynamic post-transcriptional programs, from splicing to nuclear export to transcript stability[9]. To study the dynamics of the RNA life cycle in cancer, we isolated transcript stability from other aspects of RNA regulation. We used a non-invasive method—based on 4-thiouridine labelling and capture[8,10] followed by high-throughput sequencing—to determine the decay rates for roughly 13,000 transcripts expressed by the parental MDA-MB-231 (MDA) breast cancer cell line and its *in vivo*-selected highly metastatic MDA-LM2 subline (biological quadruplicates spanning 4 time points composed of 32 total samples). A *t*-test-derived statistic (*t*-score), ranging from −1 (more stable in parental MDA) to 1 (more stable in highly metastatic MDA-LM2), was used as a measure for differential decay rates between the two cell lines. We then employed a mutual-information-based computational approach[8] to identify the *cis*-regulatory elements that may mediate the differences in transcript stability observed between the metastatic MDA-LM2 cell line and its parental MDA line. We discovered a family of structural RNA stability elements, termed sRSEs, embedded in transcripts that displayed reduced stability in highly metastatic cells (Fig. 1a and Extended Data Fig. 1a). As we show later, this broad family of GC-rich stem-loops, with a median stem length of 9 base pairs (bp) and a median loop length of 7 nucleotides, is bound by a common *trans* factor. Consistent with their higher decay rates in metastatic cells, sRSE-carrying transcripts displayed significantly reduced steady-state expression in metastatic MDA-LM2 cells relative to less metastatic MDA parental cells (Fig. 1b and Extended Data Fig. 1b). Moreover, the significantly correlated expression of these transcripts in three independent human gene-expression data sets raised the possibility of their co-regulation through a common post-transcriptional pathway mediated by this structural element (Extended Data Fig. 2a–c).

To directly assess the transcriptome-wide functionality of the identified sRSEs, we performed an in-culture cellular titration experiment in which synthetic RNA oligonucleotides harbouring tandem instances of sRSE1, the most informative representative member of the sRSE family (Extended Data Fig. 1a), were used as intracellular decoys that would bind the putative *trans* factor, preventing it from targeting endogenous transcripts[11] (Extended Data Fig. 2d). Consistent with our hypothesis, the expression levels of endogenous sRSE-carrying transcripts were significantly upregulated in cells transfected with synthetic decoys relative to their levels in cells transfected with scrambled controls (Fig. 1c and Extended Data Fig. 2e). These findings suggest that the sRSE-binding *trans* factor, when competitively titrated by the decoy, promotes transcript destabilization.

We then chose a particular sRSE, matching the motif definition of sRSE1 on a differentially destabilized transcript, for further analysis. The secondary structure of this element, determined *in silico* (M-fold[12]) and *in vitro* through differential S1/V1 nuclease digestion sequence analysis[13], matches that of the sRSE1 motif (Extended Data Fig. 3a). Additionally, mCherry reporter constructs carrying this element and its modified versions in their 3′ untranslated regions (UTRs) were used to test its functionality and establish the necessity of its underlying stem-loop structure (Extended Data Fig. 3b, c). We compared mCherry-encoding transcripts (using green fluorescent protein (GFP) as an internal control) carrying four different forms of the structural element in their 3′ UTR: an sRSE1 versus scrambled pair to reveal whether the element has a functional role in transcript stability and expression, and a structured versus unstructured mimetic pair to establish if its secondary structure is essential for its functionality (Fig. 1d). This analysis revealed that this sRSE is sufficient for suppression of expression and that its structure—not simply its sequence—is the key regulatory determinant.

We next sought to identify the sRSE-binding *trans* factor by computationally searching for candidate RNA-binding proteins (RBPs) whose expression levels correlated with sRSE-carrying transcripts across breast cancer gene expression profiles[14]. Using this approach, we identified three candidate RBPs, namely TARBP2, HEXIM1 and PPRC1, as potential post-transcriptional regulators of this regulon based on their correlated expression (Extended Data Fig. 4a). RNA interference (RNAi)-mediated

[1]Laboratory of Systems Cancer Biology, Rockefeller University, 1230 York Avenue, New York, New York 10065, USA. [2]Department of Biochemistry and Molecular Biophysics, and Department of Systems Biology, Columbia University, New York, New York 10032, USA.

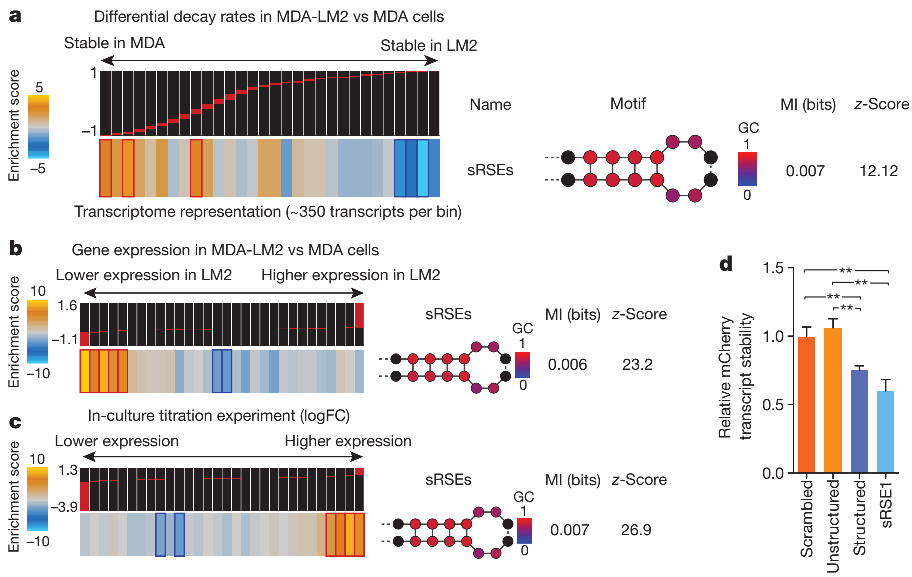


Figure 1 | A family of GC-rich structural *cis*-regulatory RNA elements are enriched in transcripts destabilized in metastatic breast cancer cells. **a**, Differential decay rates in LM2 versus MDA cells. The sRSE family of mRNA elements is significantly informative about differential transcript stability measurements between the metastatic MDA-LM2 and its parental MDA breast cancer lines. Each bin contains differential decay rate measurements for roughly 350 transcripts. From left (more stable in MDA) to right (more stable in MDA-LM2), sRSE-carrying transcripts were enriched among those destabilized in MDA-LM2 cells. In the heatmap representation of enrichment scores[8], gold entries correspond to bins with overrepresentation of sRSE-carrying transcripts, while blue bins mark underrepresentation. Statistically significant enrichments and depletions are marked with red and dark-blue borders, respectively. The sRSEs are collectively depicted as a generic stem-loop with blue and red circles marking nucleotides with low and high GC content, respectively (black positions are unconstrained regarding the number and identity of nucleotides from these positions onward; see also Methods). Also included are mutual information (MI) values and their associated *z*-scores. **b**, Gene expression in LM2 versus MDA cells. The significant enrichment of sRSE-carrying transcripts among genes with lower expression in metastatic MDA-LM2 cells is shown relative to the parental MDA cell line. Transcripts were sorted according to their $\log_2$ fold-change (logFC) of their expression levels in MDA-LM2 versus MDA cells and divided into equally populated bins from lower expression in MDA-LM2 cells (left) to higher expression (right). **c**, In-culture titration experiment. The enrichment of sRSE-carrying transcripts among the genes upregulated (based on their log-ratio) upon transfection of decoy RNA molecules harbouring engineered sRSE1 compared to scrambled controls. **d**, Transcript stability quantification for mCherry (normalized to GFP as an internal control) carrying four different forms of the sRSE, namely sRSE1, structured mimetic, unstructured mimetic and scrambled control. α-Amanitin treatments at the 0 and 8 h time points followed by total RNA extraction and complementary DNA synthesis were used to estimate relative stability between variants ($n = 6$ per sample per time point; two independent sets of biological triplicate). Error bars indicate standard error of the mean (s.e.m.). $^{**}P < 0.01$ by a one-tailed Mann–Whitney test.

knockdown followed by transcriptomic profiling revealed that silencing one of these RBPs, TARBP2, yielded a significant upregulation of sRSE-carrying transcripts (Fig. 2a and Extended Data Fig. 4b). A similar upregulation was observed upon *TARBP2* knockdown in CN-LM1a cells, an *in vivo*-selected highly metastatic breast cancer line derived from another patient's breast tumour (CN34)[6], and 293T kidney epithelial cells, suggesting a more general and physiological regulatory link between TARBP2 and sRSE-carrying transcripts (Extended Data Fig. 4c, d). Importantly, the enhanced expression of this regulon was concomitant with enhanced transcript stability upon α-amanitin treatment (Fig. 2b and Extended Data Fig. 4e). Conversely, overexpression of TARBP2 in MDA parental cells resulted in a significant downregulation of sRSE-carrying transcripts (Fig. 2c and Extended Data Fig. 4f). Moreover, the gene expression profile resulting from *TARBP2* knockdown in metastatic cells was significantly correlated to that obtained from in-culture cellular titration, consistent with the titration of TARBP2 upon sRSE-decoy RNA transfection (Extended Data Fig. 4g). TARBP2, first identified on the basis of its ability to bind the double-stranded stem portions of the human immunodeficiency virus (HIV) *tar* RNA[15], was subsequently found (through gel mobility-shift assays of synthetic RNA variants) to prefer double-stranded RNAs (for example, stem-loops) with high GC content[16]—closely resembling sRSE. TARBP2 was later found to have a physiological role in the cell by binding microRNA hairpin precursors on their path to maturation[17]. To biochemically assess direct binding of the sRSE by TARBP2, we performed ultraviolet radiation crosslinking and co-immunoprecipitation (HITS-CLIP[18]) of tagged TARBP2 and found a significant enrichment of sRSE among the TARBP2 co-immunoprecipitated mRNAs (Fig. 2d and Extended Data Fig. 5a–e). More importantly, the TARBP2-binding sites determined using HITS-CLIP enabled us to go beyond sRSE and provide a direct experimental description of the broader underlying structural elements, which we collectively termed TBSEs. TARBP2 binds both exonic and intronic TBSEs with a preference for intronic sites. TBSEs showed a high GC content along with a higher propensity to form secondary structures than expected by chance (Extended Data Fig. 5f, g). Moreover, transcript measurements from mCherry-fused sRSE/scrambled reporters showed that sRSE-dependent transcript downregulation was abrogated in the setting of *TARBP2* knockdown (Fig. 2e and Extended Data Fig. 5h). More importantly, while we initially used sRSE as a proxy to study the behaviour of the TARBP2 regulon in metastatic cells, TARBP2 HITS-CLIP provided us with the set of transcripts that are bound by TARBP2 *in vivo* that can be studied directly. These TARBP2-bound transcripts showed significant enrichment amongst the differentially destabilized transcripts in MDA-LM2 cells; they were also significantly upregulated and stabilized in the context of *TARBP2* knockdown and were downregulated in cells overexpressing TARBP2 (Extended Data Fig. 5i, j). *In vitro* binding assays using purified tagged TARBP2 and short oligo-ribonucleotides also supported our *in vivo* observations: (1) TARBP2 directly interacts with the sRSE1 element or its structured mimetic but minimally interacts with scrambled and unstructured variants (Extended Data Fig. 6a–c); (2) TARBP2-mediated co-precipitation of a large randomized RNA population followed by high-throughput sequencing resulted in an enrichment for GC-rich structured RNA variants closely resembling sRSE elements (Extended Data Fig. 6d–h). Taken together, our findings reveal that TARBP2 binds TBSEs—a family of

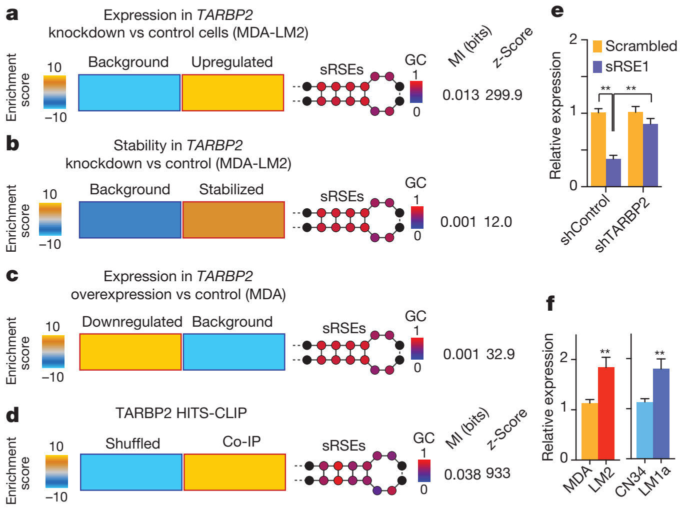


Figure 2 | TARBP2 binds and post-transcriptionally destabilizes sRSE-carrying transcripts. a, Expression of sRSE-carrying transcripts in *TARBP2* knockdown versus control (MDA-LM2). sRSE-carrying transcripts were enriched among those upregulated in the RNAi-mediated *TARBP2* knockdown in MDA-LM2 cells (compared to siControl-transfected cells). Transcripts were divided into upregulated and background groups on the basis of their expression levels in *TARBP2* knockdown cells relative to control (see Methods). b, Transcript stability in *TARBP2* knockdown versus control (MDA-LM2). The enrichment of sRSE-carrying transcripts among those whose stabilities were enhanced in *TARBP2* knockdown cells. Samples taken at the 0 and 18 h time points after α-amanitin treatment were used to estimate relative stability (see Methods). c, Expression of sRSE-carrying transcripts in *TARBP2* overexpression versus control (MDA). The sRSE regulon was enriched among transcripts downregulated in MDA-LM2 cells overexpressing *TARBP2* (relative to GFP-transfected cells). d, Significant enrichment of sRSEs among the TARBP2-binding sites (determined using HITS-CLIP). Co-IP, co-immunoprecipitation. e, The expression levels of sRSE/scrambled-fused mCherry reporters assayed in control and *TARBP2* knockdown cells ($n = 6$ per sample; two independent sets of biological triplicates). f, Relative *TARBP2* mRNA expression in MDA/MDA-LM2 and CN34/CN-LM1a cells determined using qRT–PCR ($n = 7$ per sample; three independent sets of two biological replicates and a triplicate). Error bars indicate s.e.m. **$P < 0.01$ by a one-tailed Mann–Whitney test unless otherwise specified.

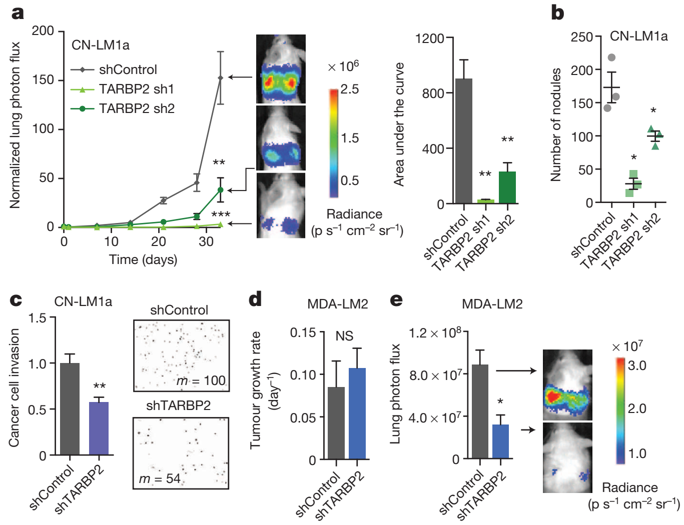


Figure 3 | Endogenous TARBP2 promotes metastatic colonization. a, Bioluminescence imaging plot of lung metastasis by CN-LM1a cells expressing short hairpin RNAs (shRNAs) targeting *TARBP2* (TARBP2 sh1 and sh2) or a control hairpin (shControl); $n = 6$ in each cohort. The area under the curve was also calculated for each mouse (change in normalized lung photon flux times days elapsed). b, The number of nodules recorded per lung section for three representative mice in each cohort ($n = 3$ per sample, extracted at day 33). $P = 0.05$ by a one-tailed Mann–Whitney test. c, Cell invasion capacity of shTARBP2 and shControl cells in the CN-LM1a background quantified using transwell invasion assays (normalized to shControl); $n = 8$ for each sample comparing shTARBP2 to shControl cells (two independent sets of biological quadruplicates). Also shown are representative images of transwell inserts along with the median count (m) for each experiment. d, Tumour growth rate for LM2 cells expressing shTARBP2 or shControl injected into the mammary fat pads of mice (day^{-1}; $n = 8$ and 6, respectively). e, Lung bioluminescence (7 days after tumour extraction) measured *in vivo* (photons s^{-1} cm^{-2} sr^{-1}; $n = 3$ and 4 for shControl and shTARBP2, respectively, in an MDA-LM2 background). Error bars indicate s.e.m. *$P < 0.05$, **$P < 0.01$ by a one-tailed Mann–Whitney test unless specified otherwise.

GC-rich apical and internal stem-loop elements within endogenous transcripts—and negatively regulates their stability and expression *in vivo* (Extended Data Fig. 6i–k).

Consistent with the observed downregulation of TARBP2-bound transcripts in metastatic MDA-LM2 cells, TARBP2 transcript and protein levels were found to be expressed at higher levels in MDA-LM2 and CN-LM1a cells relative to their parental lines (Fig. 2f and Extended Data Fig. 7a). The increased expression of TARBP2 in multiple patients' metastatic derivative lines suggested that it may have a role in metastatic progression. We also observed that primary human breast tumours that metastasized (stage IV) displayed significantly higher expression of TARBP2 relative to early stage (stages I and II) tumours, which have low metastasis rates (Extended Data Fig. 7b).

Consistent with these clinical associations, *TARBP2* knockdown significantly reduced lung metastatic colonization in CN-LM1a (~7-fold) and MDA-LM2 (~2.5-fold) cell lines as assessed by tail-vein lung colonization assays (Fig. 3a and Extended Data Fig. 7c). We also observed a significant reduction in the number of metastatic nodules in lungs of mice injected with *TARBP2* knockdown cells (Fig. 3b and Extended Data Fig. 7d). These effects on metastasis were not due to enhanced growth as the *in vitro* proliferation rate of cells was not significantly reduced upon *TARBP2* knockdown (Extended Data Fig. 7e). *TARBP2* knockdown also reduced invasiveness (~2-fold)—a phenotype required for efficient metastasis (Fig. 3c and Extended Data Fig. 7f). *TARBP2* silencing did not reduce primary tumour growth in cells growing in the mammary fat pads of mice (Fig. 3d). TARBP2 depletion did, however, significantly reduce orthotopic metastasis to the lungs (Fig. 3e and Extended Data Fig. 7g). Our findings, taken together, reveal a role for endogenous TARBP2 as a promoter of metastatic cell invasion and colonization in breast cancer.

To identify the metastasis suppressor genes that are post-transcriptionally repressed by TARBP2, we performed an unbiased analysis of the gene-expression and stability profiles of *TARBP2* knockdown cells to find transcripts that are directly bound by TARBP2 and are sensitive to modulations in TARBP2 levels. We identified four transcripts that were directly bound by TARBP2 *in vivo* (see Extended Data Fig. 8a) as well as substantially destabilized and downregulated in highly metastatic MDA-LM2 and CN-LM1a cells (validated through polymerase chain reaction with quantitative reverse transcription (qRT–PCR)), which express higher levels of TARBP2 than their isogenic parental lines (Extended Data Fig. 8b–d). These genes consisted of amyloid precursor protein (*APP*), zinc finger protein 395 (*ZNF395*), the adaptor-related protein complex 2 (*AP2A2*), and laminin β1 (*LAMB1*). Importantly, these transcripts displayed both higher stability and abundance in the context of *TARBP2* knockdown (Fig. 4a and Extended Data Fig. 8e, f). High-grade tumours, which have higher relapse rates, from three independent gene-expression compendia ($N = 821$; ref. 19) exhibited a significantly lower aggregate expression of these four genes relative to low-grade tumours (Fig. 4b). Additionally, patients whose tumours showed a reduced aggregate expression for these genes experienced significantly lower overall metastasis-free survival compared to those whose tumours expressed higher aggregate expression for these genes (Fig. 4c). These clinical associations

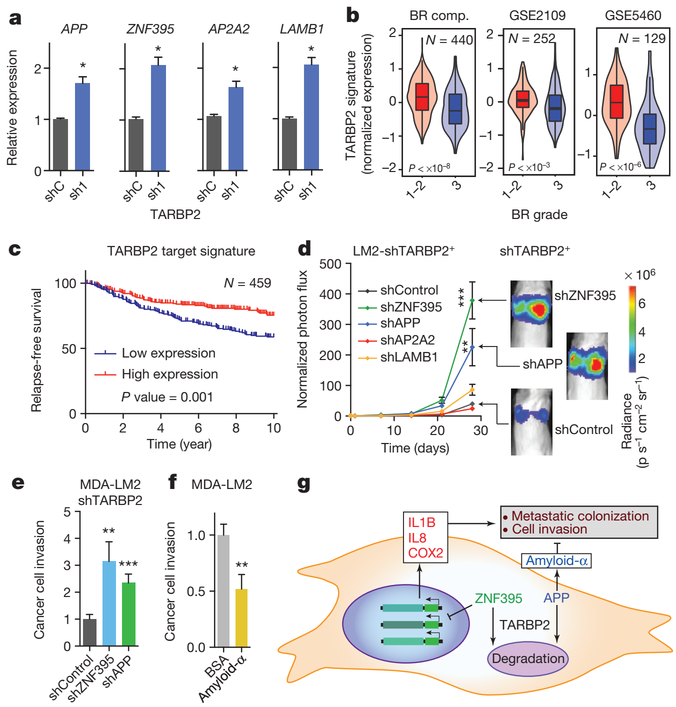


Figure 4 | TARBP2 promotes metastatic cell invasion and colonization by destabilizing *APP* and *ZNF395* transcripts. **a**, Relative expression of TARBP2 target transcripts *APP*, *ZNF395*, *AP2A2* and *LAMB1* in *TARBP2* knockdown cells versus control (MDA-LM2 background); $n = 4$ per sample (two independent sets of biological replicates). shC, shControl; sh1, TARBP2 sh1. **b**, Distribution of expression levels for the TARBP2 target signature (defined as the aggregate expression of the four targets) in tumours with Bloom–Richardson (BR) grades 1 and 2 compared to 3 in three separate data sets (BR compendium (BR comp.)[19], ExpO compendium (Gene Expression Omnibus (GEO) accession GSE2109), and GSE5460). Box plots: the bottom and the top of the box are the first and third quartile, respectively, and the central line is the median; probability density plots are superimposed. **c**, Kaplan–Meier curves for a compendium of breast cancer patients[19] ($n = 459$) showing metastasis-free survival as a function of TARBP2 target aggregate expression. *P* value is based on a Mantel–Cox log-rank test. **d**, Bioluminescence imaging plot of lung metastasis by cells expressing shRNAs targeting one of four TARBP2 targets or a control shRNA (shControl) in an MDA-LM2 shTARBP2 background; $n = 5$ in each cohort except for shZNF395, in which $n = 4$. Also shown are bioluminescence images from representative mice. **e**, Cell invasion capacity measured for *APP* and *ZNF395* knockdown cells (relative to shControl) in an MDA-LM2 shTARBP2 background; $n = 8$ for each sample (two independent sets of quadruplicates). **f**, Cell invasion capacity measured for MDA-LM2 cells with exogenously added amyloid-α relative to bovine serum albumin (BSA) as control; $n = 8$ for each sample (two independent sets of quadruplicates). **g**, Schematic of TARBP2-mediated enhancement of invasion and metastatic colonization through *APP* and *ZNF395* transcript destabilization. Error bars indicate s.e.m. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ by a one-tailed Mann–Whitney test.

and their inverse relationship to *TARBP2* transcript expression further highlight these genes' potential roles as suppressors of metastatic progression in human breast cancer and provide additional support for TARBP2 as their upstream regulator.

To identify the primary modulators of metastasis suppression among these four targets, we performed functional epistasis experiments, in which each of these genes was stably silenced in the background of *TARBP2* knockdown. In both breast cancer lines, silencing *APP* or *ZNF395* significantly enhanced metastatic colonization (Fig. 4d and Extended Data Fig. 9a–c) as well as cellular invasion by cells depleted of TARBP2 (Fig. 4e and Extended Data Fig. 9d, e). APP—previously implicated in Alzheimer's disease—is a membrane protein that is proteolytically cleaved to yield soluble products (such as, soluble amyloid-α peptide)[20], while ZNF395 is a poorly characterized transcription factor involved in the transcriptional activation of the Huntington's disease gene huntingtin[21]. Consistent with these functional findings, patients whose primary breast tumours showed reduced expression levels of *APP* and *ZNF395* experienced significantly lower survival rates than those whose tumours displayed higher levels of these genes in several independent clinical cohorts ($P < 1 \times 10^{-6}$; Extended Data Fig. 9f). These findings reveal *ZNF395* and *APP* to be downstream functional targets of TARBP2 and to associate with survival of breast cancer patients.

To determine if amyloid peptide products could directly mediate the suppressive effects of *APP* mRNA on invasion, we added soluble amyloid-α and -β to MDA-LM2 cells and quantified their invasive capacity. Interestingly, only amyoid-α treatment reduced the invasiveness of cancer cells (Fig. 4f and Extended Data Fig. 10a). Consistent with our findings, APP has also been previously implicated in ovarian cancer cell invasion[22]. Our similar observations in breast cancer metastasis suggest a more general role for *APP* as a suppressor of cancer invasion as well as a suppressor of metastatic progression by a metastatic cancer type, namely breast cancer.

On the other hand, transcriptomic profiling of *ZNF395* knockdown cells revealed upregulation of a number of established promoters of cancer progression (Extended Data Fig. 10b). Among the most highly upregulated genes, we identified and validated by qRT–PCR several known promoters of metastatic colonization and cell invasion, including *IL8* (ref. 23) (~13-fold), *IL1B* (ref. 24) (~40-fold) and *COX2* (ref. 25) (~20-fold; Extended Data Fig. 10c). Given their well-established roles in cancer progression and metastasis, we hypothesize that *ZNF395* silencing enhances metastasis in part through de-repression of these genes.

TARBP2, a component of the RISC-loading complex[17,26], has been previously implicated as a modest suppressor of tumour growth in Ewing's sarcoma through its role in microRNA maturation[27]. In contrast, our findings reveal TARBP2 to be a robust promoter of metastasis in breast cancer. Furthermore, the TARBP2-mediated post-transcriptional regulation of *APP* and *ZNF395*, and TARBP2-bound transcripts in general, proved to be independent of microRNA-mediated regulation as *DICER* knockdown did not increase transcript levels of these genes (Extended Data Fig. 10d–f). More importantly, neither *DICER* silencing nor its genetic inactivation in *Dicer*$^{-/-}$ murine cells precluded TARBP2-mediated suppression of TARBP2-bound transcripts and *APP* and *ZNF395* in particular, consistent with this network constituting a TARBP2-dependent and DICER-independent pathway (Extended Data Fig. 10g–i).

Our findings establish a novel post-transcriptional regulatory network whereby TARBP2 binding of structural hairpins contained in specific metastasis suppressor transcripts leads to their destabilization (Fig. 4g). We reveal that structural information contained in transcripts can govern cancer progression by acting as binding sites for a destabilizing RNA-binding *trans* factor. We also implicate the *TARBP2*/*APP*/*ZNF395* pathway in metastatic progression through loss-of-function, epistasis, and clinical pathological correlation analyses. Additionally, our findings reveal a microRNA-independent role for TARBP2 in gene expression regulation through mRNA destabilization. While we have focused on transcript stability, several other aspects of the RNA life cycle, such as alternative splicing and RNA localization, are impacted by structural RNA elements and may be systematically studied in the context of cancer progression using a similar approach. We speculate that small molecule or oligonucleotide-based interference with interactions between such *trans* factors and structural elements in transcripts may constitute novel routes towards therapeutic modulation of disease states.

METHODS SUMMARY

Animal experiments were conducted in accordance with protocols approved by the Institutional Animal Care and Use Committee at The Rockefeller University. See Methods for details on sequencing, computational analysis, cell culture, immunoprecipitation, western blotting, proliferation, invasion and animal studies, as well as cloning and generation of cell lines.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 September 2013; accepted 12 May 2014.

Published online 9 July 2014.

1. Hollams, E. M., Giles, K. M., Thomson, A. M. & Leedman, P. J. mRNA stability and the control of gene expression: implications for human disease. *Neurochem. Res.* **27,** 957–980 (2002).
2. Wilusz, C. J., Wormington, M. & Peltz, S. W. The cap-to-tail guide to mRNA turnover. *Nature Rev. Mol. Cell Biol.* **2,** 237–246 (2001).
3. Kumar, M. S., Lu, J., Mercer, K. L., Golub, T. R. & Jacks, T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nature Genet.* **39,** 673–677 (2007).
4. Ma, L., Teruya-Feldstein, J. & Weinberg, R. A. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* **449,** 682–688 (2007).
5. Song, S. J. *et al.* MicroRNA-antagonism regulates breast cancer stemness and metastasis via TET-family-dependent chromatin remodeling. *Cell* **154,** 311–324 (2013).
6. Tavazoie, S. F. *et al.* Endogenous human microRNAs that suppress breast cancer metastasis. *Nature* **451,** 147–152 (2008).
7. Xiao, C. & Rajewsky, K. MicroRNA control in the immune system: basic principles. *Cell* **136,** 26–36 (2009).
8. Goodarzi, H. *et al.* Systematic discovery of structural elements governing stability of mammalian messenger RNAs. *Nature* **485,** 264–268 (2012).
9. Keene, J. D. RNA regulons: coordination of post-transcriptional events. *Nature Rev. Genet.* **8,** 533–543 (2007).
10. Minn, A. J. *et al.* Genes that mediate breast cancer metastasis to lung. *Nature* **436,** 518–524 (2005).
11. Cutroneo, K. R. & Ehrlich, H. Silencing or knocking out eukaryotic gene expression by oligodeoxynucleotide decoys. *Crit. Rev. Eukaryot. Gene Expr.* **16,** 23–30 (2006).
12. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31,** 3406–3415 (2003).
13. Wan, Y., Kertesz, M., Spitale, R. C., Segal, E. & Chang, H. Y. Understanding the transcriptome through RNA structure. *Nature Rev. Genet.* **12,** 641–655 (2011).
14. Wang, Y. *et al.* Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* **365,** 671–679 (2005).
15. Gatignol, A., Buckler-White, A., Berkhout, B. & Jeang, K. T. Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. *Science* **251,** 1597–1600 (1991).
16. Gatignol, A., Buckler, C. & Jeang, K. T. Relatedness of an RNA-binding motif in human immunodeficiency virus type 1 TAR RNA-binding protein TRBP to human P1/dsI kinase and *Drosophila* staufen. *Mol. Cell. Biol.* **13,** 2193–2202 (1993).
17. Chendrimada, T. P. *et al.* TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436,** 740–744 (2005).
18. Licatalosi, D. D. *et al.* HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456,** 464–469 (2008).
19. Korpal, M. *et al.* Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization. *Nature Med.* **17,** 1101–1108 (2011).
20. Kuhn, P. H. *et al.* ADAM10 is the physiologically relevant, constitutive α-secretase of the amyloid precursor protein in primary neurons. *EMBO J.* **29,** 3020–3032 (2010).
21. Tanaka, K., Shouguchi-Miyata, J., Miyamoto, N. & Ikeda, J. E. Novel nuclear shuttle proteins, HDBP1 and HDBP2, bind to neuronal cell-specific *cis*-regulatory element in the promoter for the human Huntington's disease gene. *J. Biol. Chem.* **279,** 7275–7286 (2004).
22. Fan, X. *et al.* miR-20a promotes proliferation and invasion by targeting APP in human ovarian cancer cells. *Acta Biochim. Biophys. Sin. (Shanghai)* **42,** 318–324 (2010).
23. Strieter, R. M. *et al.* Interleukin-8. A corneal factor that induces neovascularization. *Am. J. Pathol.* **141,** 1279–1284 (1992).
24. Soria, G. *et al.* Inflammatory mediators in breast cancer: coordinated expression of TNFα & IL-1β with CCL2 & CCL5 and effects on epithelial-to-mesenchymal transition. *BMC Cancer* **11,** 130 (2011).
25. Singh, B. *et al.* COX-2 involvement in breast cancer metastasis to bone. *Oncogene* **26,** 3789–3796 (2007).
26. MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V. & Doudna, J. A. *In vitro* reconstitution of the human RISC-loading complex. *Proc. Natl Acad. Sci. USA* **105,** 512–517 (2008).
27. De Vito, C. *et al.* A TARBP2-dependent miRNA expression profile underlies cancer stem cell properties and provides candidate therapeutic reagents in Ewing sarcoma. *Cancer Cell* **21,** 807–821 (2012).

Acknowledgements We thank C. Alarcon, N. Halberg, N. Pencheva, P. Furlow and A. Nguyen for technical assistance and comments on previous versions of this manuscript. We thank H. Lee for her help with orthotopic injections. We thank P. Oikonomou for providing the reporter constructs. We thank the Darnell laboratory for their input on the HITS-CLIP methodology. We are also grateful to M. McManus and Z. Mourelatos for providing *Dicer*-null cells. We thank C. Zhao, W. Zhang and S. Dewell of the Rockefeller Genomics Resource Center for assistance with next-generation RNA-sequencing and microarray profiling. H.G. is an Anderson Cancer Center Fellow at Rockefeller University. S.F.T. is a Department of Defense Era of Hope Scholar and a Department of Defense Breast Cancer Collaborative Scholars and Innovators Award recipient. S.T. was supported by the National Human Genome Research Institute (R01 HG003219). We also acknowledge the efforts of Expression Project for Oncology (ExpO) and International Genomics Consortium (IGC) in providing public access to breast cancer gene expression profiles and annotations.

Author Contributions H.G., S.T. and S.F.T. conceived the project and designed the experiments. S.F.T. and S.T. supervised all research. H.G., S.Z., S.T. and S.F.T. wrote the manuscript. H.G., S.Z., C.G.B. and L.F. performed the experiments. S.Z. performed cancer cell invasion and reporter assays. L.F. and H.G. performed the HITS-CLIP experiment. C.G.B. and H.G. performed lung colonization assays. H.G. performed the analyses.

Author Information RNA-seq and microarray data have been deposited in the Gene Expression Omnibus under accession number GSE49649. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.F.T. (stavazoie@rockefeller.edu) or S.T. (st2744@columbia.edu).

doi:10.1038/nature13453

Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser

Christopher Kupitz[1]*, Shibom Basu[1]*, Ingo Grotjohann[1], Raimund Fromme[1], Nadia A. Zatsepin[2], Kimberly N. Rendek[1], Mark S. Hunter[1,3], Robert L. Shoeman[4], Thomas A. White[5], Dingjie Wang[2], Daniel James[2], Jay-How Yang[1], Danielle E. Cobb[1], Brenda Reeder[1], Raymond G. Sierra[6], Haiguang Liu[2], Anton Barty[5], Andrew L. Aquila[5,7], Daniel Deponte[5,8], Richard A. Kirian[2,5], Sadia Bari[9,10], Jesse J. Bergkamp[1], Kenneth R. Beyerlein[5], Michael J. Bogan[6], Carl Caleman[5,11], Tzu-Chiao Chao[1,12], Chelsie E. Conrad[1], Katherine M. Davis[13], Holger Fleckenstein[5], Lorenzo Galli[5,14], Stefan P. Hau-Riege[3], Stephan Kassemeyer[4,9], Hartawan Laksmono[6], Mengning Liang[5], Lukas Lomb[4], Stefano Marchesini[15], Andrew V. Martin[5,16], Marc Messerschmidt[8], Despina Milathianaki[8], Karol Nass[4,5,14], Alexandra Ros[1], Shatabdi Roy-Chowdhury[1], Kevin Schmidt[2], Marvin Seibert[8,17], Jan Steinbrener[4], Francesco Stellato[5], Lifen Yan[13], Chunhong Yoon[5,7], Thomas A. Moore[1], Ana L. Moore[1], Yulia Pushkar[13], Garth J. Williams[8], Sébastien Boutet[8], R. Bruce Doak[2], Uwe Weierstall[2], Matthias Frank[3], Henry N. Chapman[5,14,18], John C. H. Spence[2] & Petra Fromme[1]

Photosynthesis, a process catalysed by plants, algae and cyanobacteria converts sunlight to energy thus sustaining all higher life on Earth. Two large membrane protein complexes, photosystem I and II (PSI and PSII), act in series to catalyse the light-driven reactions in photosynthesis. PSII catalyses the light-driven water splitting process, which maintains the Earth's oxygenic atmosphere[1]. In this process, the oxygen-evolving complex (OEC) of PSII cycles through five states, S_0 to S_4, in which four electrons are sequentially extracted from the OEC in four light-driven charge-separation events. Here we describe time resolved experiments on PSII nano/microcrystals from *Thermosynechococcus elongatus* performed with the recently developed[2] technique of serial femtosecond crystallography. Structures have been determined from PSII in the dark S_1 state and after double laser excitation (putative S_3 state) at 5 and 5.5 Å resolution, respectively. The results provide evidence that PSII undergoes significant conformational changes at the electron acceptor side and at the Mn_4CaO_5 core of the OEC. These include an elongation of the metal cluster, accompanied by changes in the protein environment, which could allow for binding of the second substrate water molecule between the more distant protruding Mn (referred to as the 'dangler' Mn) and the Mn_3CaO_x cubane in the S_2 to S_3 transition, as predicted by spectroscopic and computational studies[3,4]. This work shows the great potential for time-resolved serial femtosecond crystallography for investigation of catalytic processes in biomolecules.

The first X-ray structure of PSII was determined to a resolution of 3.8 Å in 2001 (ref. 5) revealing the protein's architecture and the overall shape and location of the OEC. In 2011, Shen and co-workers achieved a breakthrough in the structural elucidation by dramatically improving crystal quality, enabling determination at 1.9 Å resolution[6]. This structure showed the OEC at near atomic resolution. However, the OEC was probably affected by X-ray damage, a fundamental problem in X-ray crystallography.

The X-ray damage problem may be overcome through the use of serial femtosecond crystallography (SFX)[2,7,8], an advance enabled by the advent of the X-ray free electron laser (XFEL). In SFX, a stream of microcrystals in their mother liquor is exposed to intense 120 Hz femtosecond XFEL pulses, thereby collecting millions of X-ray diffraction 'snapshots' in a time-frame of hours. Each X-ray FEL pulse is so intense that it destroys the sample; however, the pulse duration is so short that diffraction is observed before destruction occurs[9].

Conventional X-ray structures correspond to a temporal and spatially averaged representation of biomolecules, leading to a 'static' picture. To capture dynamic processes such as water oxidation in PSII, time-resolved X-ray data can be collected using SFX[10–12]. Conformational changes may be observed at a time-resolution ranging from femtoseconds to microseconds by combining visible laser excitation with the SFX setup and varying time delays between the optical pump and the X-ray probe snapshot. As partial reflections from crystals in random orientations are recorded, many snapshots must be collected for adequate sampling of the full reflections and three-dimensional reconstruction. A time-resolved pump-probe experiment was performed in 2010 using PSI-ferredoxin crystals as a model system, in which changes in diffraction intensities, consistent with a light-induced electron transfer process in the PSI-ferredoxin complex and dissociation of the PSI-ferredoxin complex were seen[10].

The catalytic reaction in PSII is a dynamic process. The oxygen evolution reaction is catalysed by the oxygen-evolving complex (OEC), in which the electrons are extracted from the OEC in four sequential charge separation events through the S-state cycle (Kok cycle), as shown in Fig. 1a (see ref. 1 for a review). SFX diffraction and X-ray emission spectroscopy (XES) were reported investigating the dark S_1 state and the single flash (S_2 state) of PSII[13]. The XES data show that the electronic structure of the highly radiation sensitive Mn_4CaO_5 cluster does not change during femtosecond X-ray exposure[13]. However, the quantity and quality of X-ray diffraction data was insufficient to determine if any structural changes occurred.

We report on microsecond time-resolved SFX experiments conducted at the CXI instrument[14] at the Linac Coherent Light Source (LCLS)[15]. The experimental setup is shown in Fig. 1b, c. We developed a multiple-laser illumination scheme that progressively excites the OEC in dark-adapted

[1]Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287-1604, USA. [2]Department of Physics, Arizona State University, Tempe, Arizona 85287, USA. [3]Lawrence Livermore National Laboratory, Livermore, California 94550, USA. [4]Max-Planck-Institut für medizinische Forschung, Jahnstrasse 29, 69120 Heidelberg, Germany. [5]Center for Free-Electron Laser Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany. [6]Stanford PULSE Institute, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, California 94025, USA. [7]European XFEL GmbH, Notkestrasse 85, 22607 Hamburg, Germany. [8]Linac Coherent Light Source, Stanford Linear Accelerator Center (SLAC) National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, California 94025, USA. [9]Max Planck Advanced Study Group, Center for Free-Electron Laser Science (CFEL), Notkestrasse 85, 22607 Hamburg, Germany. [10]Max-Planck-Institut für Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany. [11]Department of Physics and Astronomy, Uppsala University, Regementsvägen 1, SE-752 37 Uppsala, Sweden. [12]University of Regina, 3737 Wascana Pkwy Regina, Saskatchewan S4S 0A2, Canada. [13]Department of Physics, Purdue University, 525 Northwestern Avenue, West Lafayette, Indiana 47907, USA. [14]University of Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany. [15]Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. [16]Department ARC Centre of Excellence for Coherent X-ray Science, Department of Physics, University of Melbourne, Parkville VIC 3010, Australia. [17]Uppsala University, Sankt Olofsgatan 10B, 753 12 Uppsala, Sweden. [18]Center for Ultrafast Imaging, Luruper Chaussee 149, 22761 Hamburg, Germany.
*These authors contributed equally to this work.

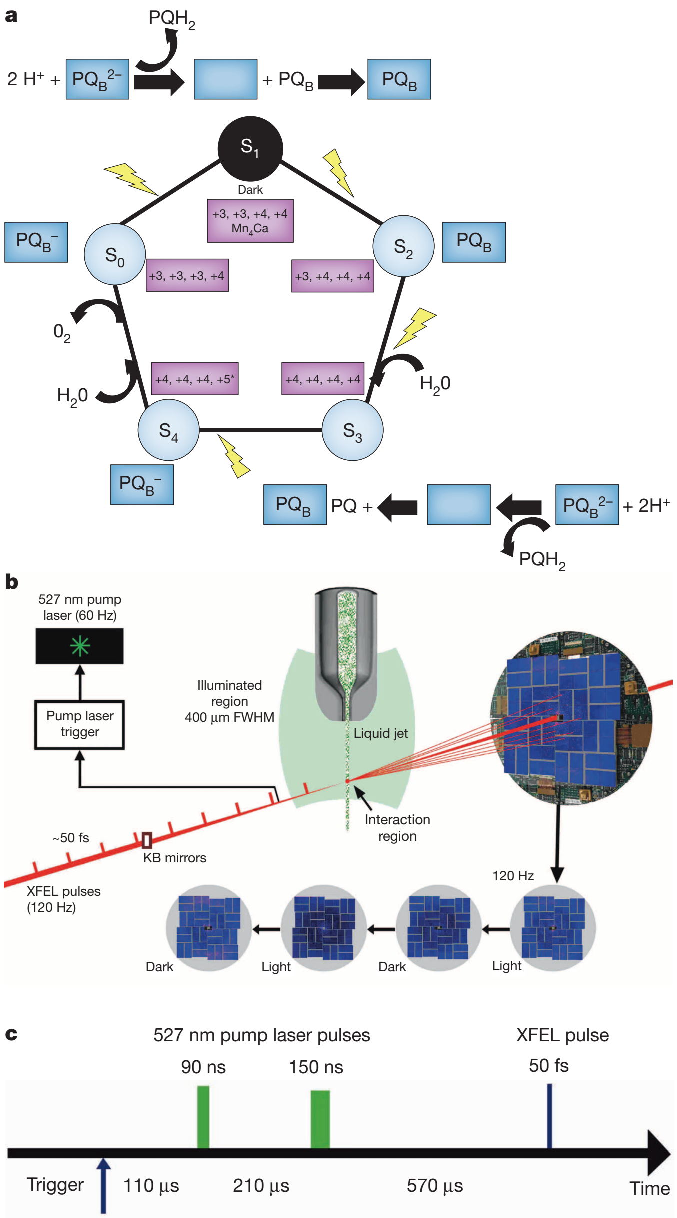


Figure 1 | Experimental schemes for the time-resolved serial femtosecond crystallography experiments on photosystem II. a, S-state scheme of the oxygen-evolving complex depicting changes the oxidation state of the 4 manganese ions of the Mn_4CaO_5 cluster in the S-state cycle (*note that the oxidation states of the Mn atoms in the S_4 state are still under debate). The scheme also indicates the reduction of the plastoquinone (PQ) to plastoquinol (PQH_2) in the Q_B site. The blank boxes represent the unoccupied PQ_B binding site. **b,** Experimental setup. The crystal-stream of photosystem II, was exposed to two subsequent optical laser pulses at 527 nm before interacting with the femtosecond X-ray FEL pulses. With a FEL frequency of 120 Hz and triggering of the laser at 60 Hz, X-ray diffraction patterns from crystals in the dark state and 'light' double-flash state alternate. **c**, Laser excitation scheme. The first 527 nm laser pulse excited the crystals 110 µs after the trigger pulse. The delay time between the first and second 527 nm laser pulse was 210 µs, with X-ray diffraction data collected 570 µs after the second laser pulse.

PSII nano/microcrystals by two laser pulses from the dark S_1 state via the S_2 state to the double-flash putative S_3 state. Not all PSII centres progress to the next S-state by a single saturating flash which could lead to heterogeneities. Therefore the S-state reached in the double-flash experiment is indicated as 'putative S_3 state' here.

The diffraction patterns collected from dark and illuminated crystals were sorted into two data sets. Using the 'hit finding' program Cheetah[16], 71,628 PSII diffraction images were identified from the dark diffraction patterns and 63,363 were identified from the double-flash patterns, see Extended Data Fig. 2 for examples. From these hits, 34,554 dark state patterns and 18,772 double-flash patterns were indexed and merged to reduce all stochastic errors using the CrystFEL software suite[17] (see Extended Data Table 2a, b). The data were indexed as orthorhombic, with unit-cell parameters of $a = 133$ Å, $b = 226$ Å, and $c = 307$ Å for the dark state, and $a = 137$ Å, $b = 228$ Å, and $c = 309$ Å for the double-flash state (for error margins see Table 1). The distributions of unit cell dimensions are shown in Extended Data Fig. 3 and Extended Data Table 2a, b. The data clearly supports an increase in unit cell dimensions in the double-flash state, with the largest difference detected for the a axis. Two factors may explain the change in unit cell constants, lower indexing rates and a slight decrease in resolution of diffraction: crystal degradation upon laser illumination or significant structural changes upon the transition from the dark state to the double-flash state, which may represent the putative S_1 to S_3 transition. To distinguish between these two possibilities, we collected data with triple-flash excitation of the PSII crystals, where at least part of the PSII centres may have reached the putative transient S_4 state. Preliminary data evaluation of the triple-flash data set (that is, putative S_4 state) shows similar unit cell dimensions and crystal quality as the dark S_1 state (see Extended Data Fig. 3 and Extended Data Table 2a). This suggests that conformational changes induced in PSII by the double-flash excitation (that is, during the putative S_1 to S_3 transition) are reversed after excitation with the third flash (in the putative S_3 to S_4 transition), as discussed in the Methods.

Diffraction data from the dark and double-flash states were evaluated to 5 Å and 5.5 Å resolution, respectively; the data refinement statistics are shown in Table 1. As each diffraction pattern represents a thin cut through reciprocal space by the Ewald sphere, only partial reflections were recorded. A high multiplicity of observations is therefore needed for each Bragg reflection to obtain full, accurate structure factors. The average multiplicity per reflection was 617 for the dark state data set and 383 for the double-flash data set over the whole resolution range (see Extended Data Tables 1a, b). Extended Data Table 2c shows a comparison of the data statistics of this work with the S_1 and S_2 data in ref. 13.

Table 1 | Statistics of the dark (S_1) and double-flash (putative S_3) data sets collected at CXI, LCLS

	Dark data set	Double-flash data set
Wavelength (Å)	2.05	2.05
Resolution range (Å)	100.6–5.0	102.3–5.5
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell length (Å)	133.3 ± 1.6, 226.3±2.1, 307.1±3.1	136.6 ± 1.5, 228.1 ± 2.3, 308.7 ± 3.8
Total reflections	28,679,554 (1,679,683)	12,476,013 (1,018,721)
Unique reflections	40,946 (2,710)	32,066 (2,651)
Multiplicity	700.35 (618.0)	388.55 (381.1)
Completeness (%)	99.98 (100)	99.88(100)
Mean I/σ (I)	10.65 (2.1)	8.03(1.75)
$CC_{1/2}$*	0.914 (0.740)	0.877 (0.635)
R_{split}	0.07 (0.37)	0.09 (0.49)
R_{work}	0.260 (0.3502)	0.280 (0.3820)
R_{free}	0.262 (0.3434)	0.290 (0.3477)
RMS† (bonds) Å	0.039	0.039
RMS† (bonds) deg	3.029	3.029
Number of atoms	49,817	49,817
Protein residues	5,214	5,214
Ramachandran favoured† (%)	97.7	97.7
Ramachandran outliers† (%)	0.2	0.2
Clashscore (Molprobity)	5.5	5.8
Mean B-factor‡ (Å^2)	33.7	33.7

$R_{split} = \sqrt{2}\dfrac{\sum_{hkl}|I_{even} - I_{odd}|}{\sum_{hkl}|I_{even} + I_{odd}|}$ See Extended Data Fig. 4 for a comparison of R_{split} vs resolution.

Numbers in parentheses refer to values for the highest resolution shell.

* $CC_{1/2}$ is Pearson's coefficient calculated as described in ref. 29.

† The values for the RMS for bonds and angles as well as the Ramachandran values are positively biased by the high resolution model with PDB accession code 3ARC.

‡ The B-factors were taken from the high resolution model with PDB accession code 3ARC and not refined.

The data were phased by molecular replacement using a truncated version of the 1.9 Å structure (PDB accession code 3ARC)[6]. Rigid body refinement (phenix.refine) was performed for both the dark and double-flash structures (see Methods for further details on molecular replacement and refinement). To reduce model bias, we calculated omit maps and simulated annealed maps (SA-omit maps) for the dark and double-flash data, deleting the coordinates of the Mn_4CaO_5 cluster from the model.

Figure 2a–c shows the arrangement of protein subunits and cofactors of photosystem II, including the electron transport chain. The comparison of the electron density maps for the dark state (green) and the double-flash state (white) at a contour level of 1.5 σ is shown in Fig. 2d–f. Both maps show clear electron densities for the transmembrane helices as well as loops and cofactors. Additional electron density maps for representative structural elements of PSII are shown in Extended Data Figs 5, 6, 7 and 8. Overall, the protein fits into the electron densities for the dark and double-flash states and matches with the high resolution structural model. However, differences appear in regions of the Mn_4CaO_5 cluster and the acceptor side, where the quinones and the non-haem iron are located. Determining the significance of these changes and their correlations is complicated due to the resolution limit of the data. Figure 2g–i shows detailed views of the loops at the acceptor side of PSII. The quinones are not visible at the current resolution of 5 Å. The maps indicate differences between the electron densities of the dark and double-flash states in the loop regions and also in the position of the non-haem iron that is coordinated by the loops.

We now focus on the structure in the undamaged dark S_1 state of the metal cluster in the OEC and the potential light-induced structural changes that may occur during the S-state transition. Extended Data Fig. 8 shows the SA-omit map of the OEC in the dark S_1 state for the Mn cluster in PSII with the 1.9 Å X-ray structure in ref. 6. Interestingly, the electron-density map of the dangler Mn atom from the 1.9 Å structure is located outside the dark S_1 state electron density, a feature also visible in the electron density map of ref. 13. These structural observations are consistent with spectroscopic results, which indicate that the distance between the dangler Mn and the Mn_3O_xCa distorted cubane is indeed shorter in the dark S_1 state than in the 1.9 Å structure based on the synchrotron data, which might be influenced by X-ray induced reduction of the Mn ions in the metal cluster[18,19]. This shorter distance is in agreement with density function theory (DFT) studies[4,18,20] based on the 1.9 Å structure of PSII[6], however, the current resolution limit of 5 Å does not allow a quantitative assessment.

The mechanism of water splitting is intensely debated and many models have been proposed. The recent 1.9 Å X-ray structure[6] formed the basis

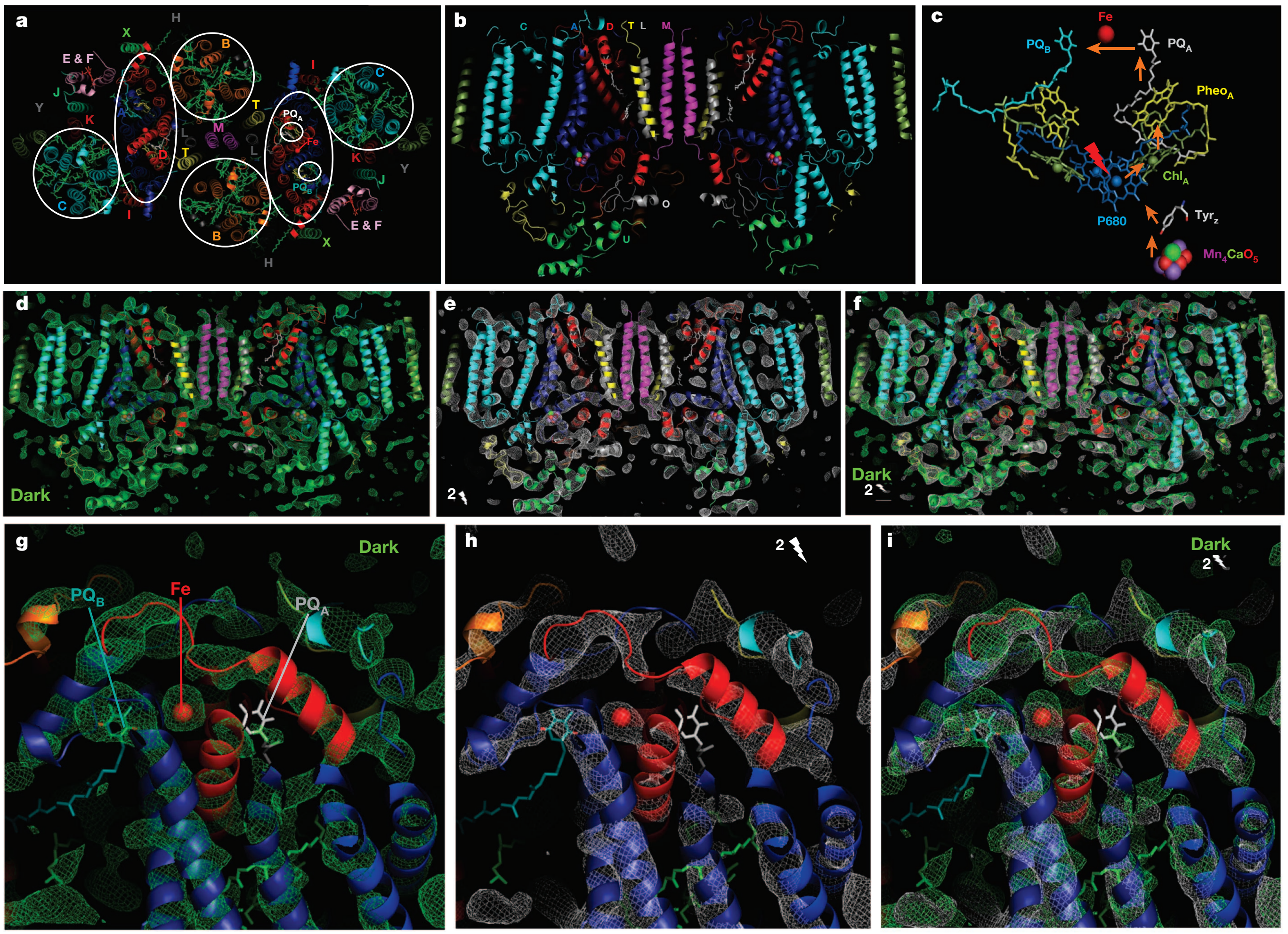


Figure 2 | Overall structure and omit map electron density of photosystem II. **a,** Transmembrane helices and cofactors in photosystem II (stromal view density map). The proteins are named according to their genes and labelled with coloured letters. **b,** Side view of PSII at its longest axis along the membrane plane. **c,** Electron transport chain of PSII (P680 (blue), accessory chlorophylls (olive-green), pheophytins (yellow) and plastoquinones PQ_A (white) and PQ_B (cyan)); atoms of the OEC are depicted as spheres (Mn purple, Ca green, O light red). **d–f,** Omit map electron densities (view as in **b**) at 1.5 σ for the dark state (S_1) (green) (**d**), double-flash state (putative S_3 state) (white) (**e**) and overlay of the two omit maps (**f**). **g–i,** Omit maps (1.5 σ) of the electron acceptor side of photosystem II for the dark (S_1) (green) (**g**), double-flash (putative S_3 state) (white) (**h**) and overlay of the two omit maps (**i**). Note that changes include a shift of the electron density of the non-haem iron.

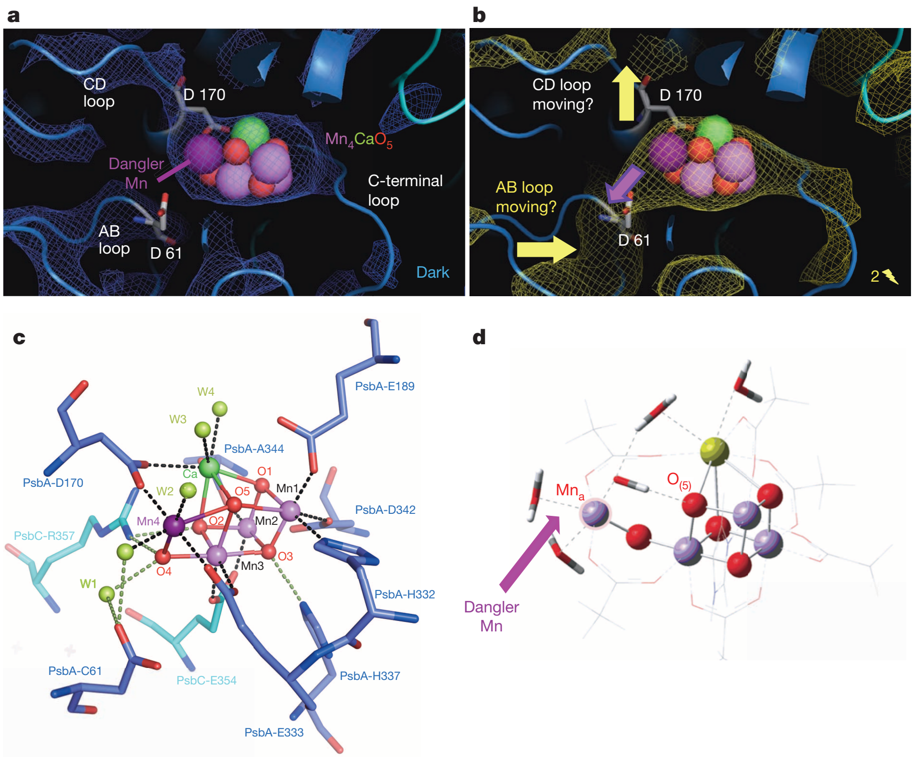


Figure 3 | The OEC simulated annealed omit maps. **a**, **b**, At 1.5 σ for dark and double-flash states of the Mn_4CaO_5 cluster of PSII for the dark S_1-state (blue) (**a**) and double-flash, putative S_3 state (**b**) with the 1.9 Å structural model (3ARC) from ref. 6. Mn in the distorted Mn_3O_xCa cubane (Mn-1 to Mn-3) (light-pink), dangler manganese (Mn-4) (violet), calcium (green) and oxygen (red). **c**, 1.9 Å crystal structure of the Mn_4CaO_5 cluster with ligands from ref. 6 (PDB accession code: 3ARC). **d**, Proposed model of the S_3 state based on DFT calculations by Isobe *et al.*[4] (reproduced with permission of The Royal Society of Chemistry). Larger diversions in the SA-omit map of the double-flash (putative S_3 state) include potential movement of the loop connecting transmembrane helices C/D (CD loop) with D170 and AB loop (with D61), and an increase of the distance between the dangler Mn and the Mn_3O_xCa cubane (violet arrow).

for more detailed theoretical studies of the process, yet the proposed mechanisms differ[4,20–22]. Based on our time-resolved SFX (TR-SFX) structural data, we looked for differences between the electron-density maps of the OEC, derived from the dark and the double-flash data sets. Figure 3a, b shows the SA-omit maps calculated for dark (blue) and double-flash state (yellow) and compared with the model of the metal cluster from the 1.9 Å structure[6] (Fig. 3c). The Mn_4CaO_5 cluster was omitted from the model for the calculation of the SA-omit map, which is based on annealing at a virtual temperature of 5000 K to minimize phase bias. The SA-omit electron densities of the dark and double-flash states differ in the shape and position, as well as in the protein environment, of the Mn_4CaO_5 cluster. The dark state simulated-annealed (SA)-omit electron density for the OEC protein environment matches the model of the 1.9 Å structure[6], whereas the SA-omit map of the double-flash state differs significantly. Any interpretation of changes in the protein environment of the OEC is highly speculative at a resolution of 5 Å and heterogeneities in the S-state transitions. However, the SA-omit map of the double-flash state is suggestive of conformational changes which may indicate a movement of the CD loop (including the ligand D170) away from the cluster. If confirmed at higher resolution, this could explain mutagenesis studies that raised questions about the ligand-role of D170 in the higher S-states[23]. Furthermore, in the double-flash state, the electron density of the metal cluster extends and shows a new connection to the AB loop at the site where D61 is located. Although D61 only serves as a second sphere ligand in the 1.9 Å crystal structure[6], mutagenesis studies indicated an important role in the water oxidation process, as the S_2 to S_3 transition is blocked in D61 mutants.

The change in the electron-density of the OEC is suggestive of an increase in the distance between the cubane and the dangler Mn and distortion in the cubane in the double-flash state. The observed electron densities (Fig. 3a, b) of the dark state and double-flash state are consistent with conformational changes predicted in a recent DFT study of the S_3 state in ref. 4, shown in Fig. 3d. The increased distance between the cubane and dangler Mn could allow the second 'substrate' water molecule to bind between the Mn_3O_xCa cubane and the dangler Mn in the S_2 to S_3 state transition. It was shown by extended X-ray absorption fine structure (EXFAS) spectroscopy that the Mn–Ca^{2+} distances in the Mn_3O_xCa cubane shrink in the S_3 state[24]. Although the Jahn-Teller effect extends the distances between metals in the lower S-states of the OEC (Mn oxidation states +II, +III and +IV), a shrinking of the Mn_3O_xCa cubane is predicted in the S_3 state when all 4 Mn in the OEC have reached the oxidation state +IV. A comparison of the electron density in the dark and the double-flash states may indeed suggest an overall decrease in the dimension of the Mn_3O_xCa cubane in the double-flash state, which is in good agreement with the proposed S_3 state EXAFS and XES models[25] (more detail provided in the Supplementary Discussion). The consistency of spectroscopy and DFT studies with our observations may provide preliminary indications that a significant fraction of the OEC centres in our crystals have reached the S_3 state in the double-flash experiment.

Our time-resolved SFX study captures the image of PSII after it has been excited by 2 saturating flashes and provides experimental evidence for structural changes occurring in the putative S_3 state of the OEC, accompanied by structural changes at the acceptor side of PSII. As the resolution is limited to 5 Å, the interpretation of the changes observed is preliminary. This work is a proof-of-principle that time-resolved SFX can unravel conformational changes at moderate resolution and may lay the foundation for solving high resolution structures of PSII at all stages of the water oxidation process in the future. To unlock the secrets of the water-splitting mechanism by TR-SFX at atomic detail, the resolution must be further improved and structures must be determined from all the S-states with multiple time delays.

METHODS SUMMARY

Here, we describe microsecond time resolution optical pump/X-ray probe SFX experiments on PSII nano/microcrystals, to study conformational changes in PSII in the transition from the dark to the double-flash state of PSII, where structures were determined at 5 and 5.5 Å resolution, respectively. Nanocrystal growth for SFX was performed using a free-interface diffusion technique (see Extended Data Fig. 1a–e). The size and crystallinity of the samples were monitored by dynamic light scattering (DLS) and second order non-linear imaging of chiral crystals (SONICC)[26]. Time-resolved SFX data were collected from PSII crystals delivered to the X-ray FEL interaction region at room temperature in a liquid jet[27], The crystals were progressed along the S-state cycle[28] from the dark S_1 to the putative S_3 state by two saturating laser flashes before the structure was probed by interaction with the X-rays flashes (see Fig. 1b, c and Methods for details). The structure factors and coordinates have

been deposited in the Protein Data Bank and accession codes for S_1 and putative S_3 states are 4PBU and 4Q54 respectively.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 November 2013; accepted 4 May 2014.

Published online 9 July; corrected online 10 September 2014 (see full-text HTML version for details).

1. Renger, G. Mechanism of light induced water splitting in photosystem II of oxygen evolving photosynthetic organisms. *Biochim. Biophys. Acta* **1817,** 1164–1176 (2012).
2. Chapman, H. N. *et al.* Femtosecond X-ray protein nanocrystallography. *Nature* **470,** 73–77 (2011).
3. Navarro, M. P. *et al.* Ammonia binding to the oxygen-evolving complex of photosystem II identifies the solvent-exchangeable oxygen bridge (μ-oxo) of the manganese tetramer. *Proc. Natl Acad. Sci. USA* **110,** 15561–15566 (2013).
4. Isobe, H. *et al.* Theoretical illumination of water-inserted structures of the $CaMn_4O_5$ cluster in the S_2 and S_3 states of oxygen-evolving complex of photosystem II: full geometry optimizations by B3LYP hybrid density functional. *Dalton Trans.* **41,** 13727–13740 http://dx.doi.org/10.1039/c2dt31420g (2012).
5. Zouni, A. *et al.* Crystal structure of photosystem II from *Synechococcus elongatus* at 3.8 Å resolution. *Nature* **409,** 739–743 (2001).
6. Umena, Y., Kawakami, K., Shen, J. R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473,** 55–60 (2011).
7. Boutet, S. *et al.* High-resolution protein structure determination by serial femtosecond crystallography. *Science* **337,** 362–364 (2012).
8. Redecke, L. *et al.* Natively inhibited *Trypanosoma brucei* cathepsin B structure determined by using an X-ray laser. *Science* **339,** 227–230 (2013).
9. Barty, A. *et al.* Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. *Nature Photon.* **6,** 35–40 (2012).
10. Aquila, A. *et al.* Time-resolved protein nanocrystallography using an X-ray free-electron laser. *Opt. Express* **20,** 2706–2716 (2012).
11. Neutze, R. & Moffat, K. Time-resolved structural studies at synchrotrons and X-ray free electron lasers: opportunities and challenges. *Curr. Opin. Struct. Biol.* **22,** 651–659 (2012).
12. Spence, J. C. H., Weierstall, U. & Chapman, H. N. X-ray lasers for structural and dynamic biology. *Rep. Prog. Phys.* **75,** 102601 (2012).
13. Kern, J. *et al.* Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature. *Science* **340,** 491–495 (2013).
14. Boutet, S. & Williams, G. J. The coherent X-ray imaging (CXI) instrument at the Linac Coherent Light Source (LCLS). *New J. Phys.* **12,** 035024 (2010).
15. Emma, P. *et al.* First lasing and operation of an angstrom-wavelength free-electron laser. *Nature Photon.* **4,** 641–647 (2010).
16. Barty, A. *et al.* Cheetah: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J. Appl. Crystallogr.* **47,** 1118–1131 (2014).
17. White, T. A. *et al.* CrystFEL: a software suite for snapshot serial crystallography. *J. Appl. Crystallogr.* **45,** 335–341 (2012).
18. Luber, S. *et al.* S1-state model of the O_2-evolving complex of photosystem II. *Biochemistry* **50,** 6308–6311 (2011).
19. Davis, K. M., Kosheleva, I., Henning, R. W., Seidler, G. T. & Pushkar, Y. Kinetic modeling of the X-ray-induced damage to a metalloprotein. *J. Phys. Chem. B* **117,** 9161–9169 (2013).
20. Ames, W. *et al.* Theoretical evaluation of structural models of the S2 state in the oxygen evolving complex of photosystem II: protonation states and magnetic interactions. *J. Am. Chem. Soc.* **133,** 19743–19757 (2011).
21. Siegbahn, P. E. Water oxidation mechanism in photosystem II, including oxidations, proton release pathways, O–O bond formation and O_2 release. *Biochim. Biophys. Acta* **1827,** 1003–1019 (2013).
22. Rivalta, I., Brudvig, G. W. & Batista, V. S. Oxomanganese complexes for natural and artificial photosynthesis. *Curr. Opin. Chem. Biol.* **16,** 11–18 (2012).
23. Debus, R. J., Strickler, M. A., Walker, L. M. & Hillier, W. No evidence from FTIR difference spectroscopy that aspartate-170 of the D1 polypeptide ligates a manganese ion that undergoes oxidation during the S_0 to S_1, S_1 to S_2, or S_2 to S_3 transitions in photosystem II. *Biochemistry* **44,** 1367–1374 (2005).
24. Pushkar, Y., Yano, J., Sauer, K., Boussac, A. & Yachandra, V. K. Structural changes in the Mn_4Ca cluster and the mechanism of photosynthetic water splitting. *Proc. Natl Acad. Sci. USA* **105,** 1879–1884 (2008).
25. Dau, H., Zaharieva, I. & Haumann, M. Recent developments in research on water oxidation by photosystem II. *Curr. Opin. Chem. Biol.* **16,** 3–10 (2012).
26. Wampler, R. D. *et al.* Selective detection of protein crystals by second harmonic microscopy. *J. Am. Chem. Soc.* **130,** 14076–14077 (2008).
27. Weierstall, U., Spence, J. C. H. & Doak, R. B. Injector for scattering measurements on fully solvated biospecies. *Rev. Sci. Instrum.* **83,** 035108 (2012).
28. Dekker, J. P., Plijter, J. J., Ouwehand, L. & Vangorkom, H. J. Kinetics of manganese redox transitions in the oxygen-evolving apparatus of photosynthesis. *Biochim. Biophys. Acta* **767,** 176–179 (1984).
29. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336,** 1030–1033 (2012).
30. Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. & Echols, N. New Python-based methods for data processing. *Acta Crystallogr. D* **69,** 1274–1282 (2013).
31. Kok, B., Forbush, B. & McGloin, M. Cooperation of charges in photosynthetic O_2 evolution-I. A linear four step mechanism. *Photochem. Photobiol.* **11,** 457–475 (1970).

Supplementary Information is available in the online version of the paper.

Acknowledgements Experiments were carried out at the Linac Coherent Light Source (LCLS), a national user facility operated by Stanford University on behalf of the US Department of Energy (DOE), Office of Basic Energy Sciences (OBES). This work was supported by the following agencies: the Center for Bio-Inspired Solar Fuel Production, an Energy Frontier Research Center funded by the DOE, Office of Basic Energy Sciences (award DE-SC0001016), the National Institutes of Health (award 1R01GM095583), the US National Science Foundation (award MCB-1021557 and MCB-1120997), the DFG Clusters of Excellence 'Inflammation at Interfaces' (EXC 306) and the 'Center for Ultrafast Imaging'; the Deutsche Forschungsgemeinschaft (DFG); the Max Planck Society, the Atomic, Molecular and Optical Sciences Program; Chemical Sciences Geosciences and Biosciences Division, DOE OBES (M.J.B.) and the SLAC LDRD program (M.J.B., H.L.); the US DOE through Lawrence Livermore National Laboratory under the contract DE-AC52-07NA27344 and supported by the UCOP Lab Fee Program (award no. 118036) and the LLNL LDRD program (12-ERD-031); the Hamburg Ministry of Science and Research and Joachim Herz Stiftung as part of the Hamburg Initiative for Excellence in Research. The research at Purdue University was supported by the U.S. Department of Energy (DOE), Office of Basic Energy Sciences DE-FG02-12ER16340 (Y.P.) and the National Science Foundation Graduate Research Fellowship under Grant 0833366 (K.M.D.). We also want to thank the National Science Foundation for providing funding for the publication of this work through the BioFEL Science Technology Center (award 1231306). We thank H. Isobe, M. Shoji, S. Yamanaka, Y. Umena, K. Kawakami, N. Kamiya, J. R. Shen and K. Yamaguchi for permission to show a section of Fig. 6 of their publication ref. 4 in Fig. 3d of this publication. We thank R. Neutze and his team for support and discussions during joint beamtime for the PSII project and his projects on time-resolved wide-angle scattering studies. We thank A. T. Brunger for discussions concerning data analysis. We thank T. Terwilliger for support with parameter setting of phenix.autobuild program for the SA-omit maps. We also wish to thank R. Burnap for discussions concerning interpretation of results of ligand mutagenesis. We thank J. D. Zook for his contributions concerning plastoquinone quantification. We thank M. Zhu for helping to create high resolution figures for this publication. We thank Raytheon for support of our studies by providing night-vision devices.

Author Contributions C.K., I.G., R.F., M.S.H., R.L.S., A.R., K.S., G.J.W., S. Boutet, H.N.C., U.W., R.B.D., M.F., J.C.H.S. and P.F. contributed to the design of the experiment; C.K., I.G., K.N.R., J.-H.Y., D.E.C., B.R., C.E.C. and S.R.-C. worked on cell growth and photosystem II isolation; J.J.B., T.A.M. and A.L.M. worked on plastoquinone synthesis; C.K., I.G., K.N.R., D.E.C., B.R. and J.J.B. worked on biochemical and biophysical characterization of the photosystem II samples; C.K., K.M.D., L.Y. and Y.P. worked on EPR experiments to confirm the S_3 population; C.K., I.G., M.S.H., D.E.C. and P.F. developed nano/microcrystallization conditions of photosystem II; C.K., I.G., R.F., K.N.R., M.S.H. and D.E.C. grew crystals of photosystem II; C.K., I.G., R.F., K.N.R., J.-H.Y., D.E.C., R.G.S., H. Laksmono, M.J.B., T.-C.C. and P.F. conducted biophysical characterization of photosystem II crystals; C.K., I.G., L.G., M.L., L.L., J. Steinbrener, F.S. and P.F. designed and/or fabricated calibration or backup samples; C.K., I.G., D.W., D.J., D.D., U.W., R.B.D. and P.F. tested and optimized buffer and crystal suspension conditions for injection; D.W., D.J., D.D., R.A.K., U.W. and R.B.D. designed and produced nozzles; R.B.D., U.W., R.L.S., D.W., D.J., D.D., R.A.K., S. Bari. and L.L. developed and operated the injector; R.L.S., J. Steinbrener and L.L. developed and operated the sample delivery system and the anti-settling device; S. Boutet, M.M. and G.J.W. developed diffraction instrumentation; M.M., M.S., G.J.W. and S. Boutet set up and operated the CXI beamline; M.S.H., R.A.K., D.M., S. Boutet, M.F. and P.F. designed and optimized the laser excitation scheme and aligned the lasers; C.K., S. Basu., I.G., R.F., N.A.Z., M.S.H., R.L.S., T.A.W., D.W., D.J., D.E.C., H.F., H. Laskmono, H. Liu, A.B., A.L.A., D.D., R.A.K., S. Bari., K.R.B., M.J.B., T.-C.C., L.G., S.K., C.C., M.L., M.M., K.N., M.S., J. Steinbrener, F.S., C.Y., G.J.W., S. Boutet, H.N.C., U.W., R.B.D., M.F., J.C.H.S. and P.F. collected X-ray diffraction data at the CXI beamline; S. Basu, R.F., N.A.Z., T.A.W., H. Liu, A.B., A.L.A., R.A.K., K.R.B., S.K., K.N., L.G., C.Y., J.C.H.S. and P.F. analysed the femtosecond crystallography X-ray diffraction data; T.A.W., A.B., A.L.A., R.A.K. and H.N.C. developed the data evaluation and/or hit finding programs; S. Basu, R.F. and N.A.Z. merged the 3D data; S. Basu and R.F. refined the structure and calculated the electron density maps; S. Basu, R.F., N.A.Z. and P.F. designed and made the figures; R.L.S., T.A.W., D.W., D.J., R.L.S., A.B., A.L.A., A.R., K.S., S.M., A.V.M., S.P.H.-R., R.G.S., H.N.C., U.W., R.B.D., M.F., J.C.H.S., T.A.M. and A.L.M. contributed to the writing of the manuscript with discussion, comments or edits; C.K., S. Basu, R.F., N.A.Z., K.N.R., H.N.C., M.F., J.C.H.S. and P.F. contributed to the interpretation of the results; C.K., S. Basu, I.G., R.F., N.A.Z., K.N.R., C.E.C., H.N.C., U.W., R.B.D., M.F., S.R.-C., J.C.H.S. and P.F. wrote and edited the manuscript with discussion and input from all authors.

Author Information The structure factors and coordinates have been deposited in the Protein Data Bank and accession codes for S_1 and putative S_3 states are 4PBU and 4Q54, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.F. (pfromme@asu.edu).

CORRIGENDUM

doi:10.1038/nature13550

Corrigendum: Mitofusin 2 tethers endoplasmic reticulum to mitochondria

Olga Martins de Brito & Luca Scorrano

Nature **456,** 605–610 (2008); doi:10.1038/nature07534

In Fig. 1a of this Article, the representative image of a volume-rendered three-dimensional reconstruction of a z-stack of confocal images of endoplasmic-reticulum-targeted yellow fluorescent protein (ER-YFP) in a *Mfn2*$^{-/-}$ cell expressing MFN2^{IYFFT} and that of a *Mfn1*$^{-/-}$ cell appear to be duplicated. Because the original raw data could not be located, we were unable to verify definitively whether the data in the original figure were indeed inadvertently duplicated. We therefore obtained new images under the same experimental conditions. The correct images of representative volume-rendered three-dimensional reconstruction of z-stacks of confocal images of ER-YFP in *Mfn1*$^{-/-}$ cells, *Mfn2*$^{-/-}$ cells and *Mfn2*$^{-/-}$ cells expressing MFN2^{IYFFT} are shown in Fig. 1 of this Corrigendum. This does not affect any of our results.

Correspondence should be addressed to L.S. (luca.scorrano@unipd.it).

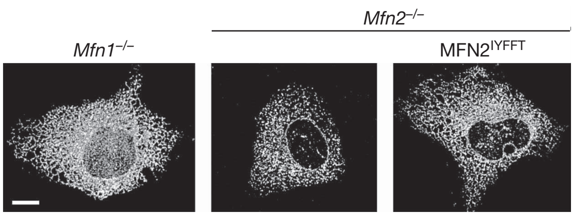


Figure 1 | This figure shows the results of the repeated experiments of Fig. 1a of this Article. Three-dimensional reconstructions of endoplasmic reticulum in mouse embryonic fibroblasts of the indicated genotype co-transfected with ER-YFP and the specified plasmids. Scale bar, 10 μm.

CORRIGENDUM

doi:10.1038/nature13631

Corrigendum: Direct recording and molecular identification of the calcium channel of primary cilia

Paul G. DeCaen, Markus Delling, Thuy N. Vien & David E. Clapham

Nature **504,** 315–318 (2013); doi:10.1038/nature12832

In this Letter, Fig. 1c and e contained errors, which are corrected in Fig. 1 of this Corrigendum. In the key of Fig. 1c, the extracellular conditions for the Ba^{2+} and NMDG currents were wrongly listed: the purple trace should be $BaCl_2$ (not RbCl) and the grey trace should be NMDG (not $BaCl_2$). This correction accurately reflects the selectivity of the channel as stated on page 315, which should have cited Fig.1c, rather than Fig. 1d, as follows: "The outwardly rectifying current was cation-non-selective (Fig. 1c) with relative permeabilities of $Ca^{2+} \approx Ba^{2+} > Na^{+} \approx K^{+} >$ NMDG". In Fig. 1e, the channel density for the primary cilia (red bar labelled 'PKD1L1/2L1') was wrong, owing to an incorrect estimation of open probability for the channel (0.015 instead of 0.067). This correction reduces the estimated channel density to a value (29 channels per μm^2) that is about 4.5 times smaller than the value we initially reported in Fig. 1 and on page 315 (128 channels per μm^2). This error does not alter the key findings of Fig. 1e (that the cilia are densely populated with PKD1L1/2L1 channels) or the main conclusions of the Letter.

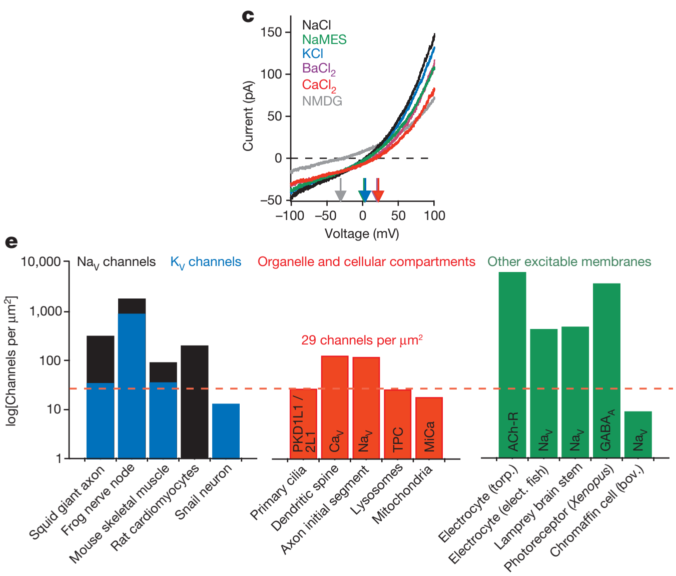


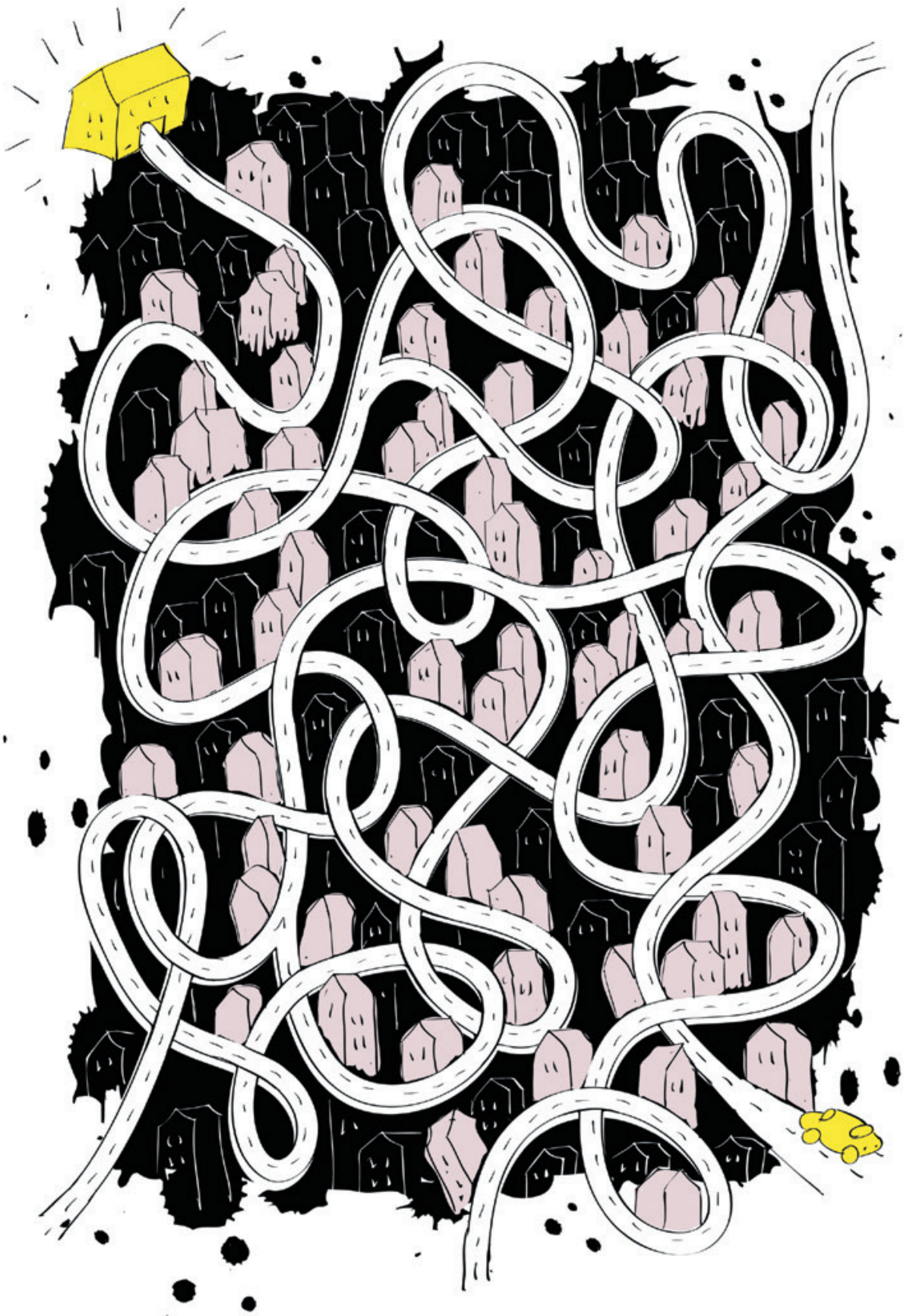
Figure 1 | This figure shows the corrected Fig. 1c and e of the original Letter.

CAREERS

@NATUREJOBS Follow us on Twitter for the latest news and features **go.nature.com/e492gf**

NATUREJOBS BLOG Up-to-the-minute careers news and tips **go.nature.com/ielkkf**

NATUREJOBS For the latest career listings and advice **www.naturejobs.com**



VOOK/SHUTTERSTOCK

JOB SATISFACTION

Divided opinions

Financial woes are marring researchers' enjoyment of their work.

BY PAUL SMAGLIK

Researchers around the world love their work, but tight funding is eroding their spirits, according to this year's *Nature* Careers salary survey. Although nearly two-thirds of the survey's 7,216 respondents across the world report being satisfied or very satisfied with their job, nearly half say that they think that the morale in their lab or department is slipping, and two-thirds of those who responded to the question do not believe that the funding environment is improving (see 'Money and morale').

The survey asked participants not only about morale in their lab or department but also about the level and accessibility of science funding. It also asked people where their funding came from, such as government grants or contracts, private grants, royalties or venture-capital funds. Participants could also rate how 15 factors — including salary, benefits, financial resources, interest in their work and availability of funding — affect their job satisfaction. And they were asked to rate their level of job satisfaction and indicate whether that had changed in the past year.

Nearly half of respondents across all participating nations say that the availability of funding is cutting into their job satisfaction. That was the biggest negative indicator of job satisfaction in the survey; more than salary, interest in their work and level of guidance (see 'When guidance is important').

Two in every five people also said that the availability of financial resources — of their nation, institution, department or supervisor — negatively affected their satisfaction with their job. Participants from several nations, including the United States, the United Kingdom, Japan and Spain, say that it was more difficult to secure funding last year than the year before.

Respondents are, at least, engaged by their work. Around four-fifths counted interest in their jobs as a positive factor, the most for any indicator of satisfaction. Almost two-thirds say that they are satisfied with their level of independence while more than half said that their colleagues had positively affected their job satisfaction and that they are happy with the location of their workplace. Salary had the most mixed results — roughly equal numbers rated it as positive, negative and neutral in terms of how it affected their job satisfaction.

Job satisfaction seems to rise with age. Three in five of those aged 25–54 are very satisfied ▶

or satisfied. But at age 55 and up, those numbers increase: three in four of those 55–64 say that they are satisfied or very satisfied, and more than four in five of those 65 and older say that they are satisfied or very satisfied. Unsurprisingly, some of the respondents in this age group say that they do not worry as much as their younger colleagues about winning grants.

DIM VIEW

We interviewed some of the respondents after the survey — and many of them said that budgetary problems in their country, which they have faced since the global financial crisis of 2008, threaten their long-term satisfaction. Many also said that they do not see any quick turnaround in the dim situation for science funding.

Interviewed respondents also said that the funding malaise is starting to affect multiple aspects of their job satisfaction. Some said that worries about funding caused them to downgrade their job outlook from very satisfied to satisfied; that spending more time writing grants means less time for research; and that it has created uncertainty or is making the transition to their next career stage more challenging.

Several postdoc respondents, for example, said that they know that they will need to land a grant to kick-start their future, and that they are beginning to become more aware of the proliferation of less-permanent positions owing to budget constraints. "I have a very unstable position so I cannot develop all the things I would like to do," says Victor Ladero, a postdoc at the Dairy Institute of Asturias in Villaviciosa, Spain. "I cannot plan for the long, even the middle, term." Because of these limitations, he says that he feels "neutral" in terms of job satisfaction.

Garry Buettner, a radiation oncologist at the University of Iowa in Iowa City, is concerned that this financially constrained environment will discourage talented people from becoming scientists. "Where are the opportunities?" he asks. "We are supposed to be training our replacements. But where will they go? Where is our investment in the future?" He also feels responsible for younger scientists working with him. "They are vulnerable to changes in funding," he says. "This is what keeps me up at night."

In discussing how the lack of financial resources has diminished their job satisfaction, several people noted that rising funding and budget pressures are not the only problem — research costs have increased too. And they said that their universities are relying more on researcher grants to cover operating costs, which leaves less for the researchers. "If I get a grant for US$100,000, the university gets half," says Buettner. Not long ago, he could spend most of his grant on personnel and direct research costs.

> ***"Everyone would always like more money for research."***

One in five respondents strongly agreed that it was more difficult to secure funding in 2013 than in previous years, while another one-third said that it remained the same as before. Scott Steppan, a geneticist at Florida State University in Tallahassee, notes that faculty scientists now need to write more grant applications if they hope to maintain their level of funding. Changes in the review process — made in part to accommodate the increase in applications and decrease in reviewers — are exacerbating the problem, he says.

Paul Roepe, a chemist at Georgetown University in Washington DC, said that the grant-review process seems more "arbitrary" now, since many quality projects do not get funded because of increased competition for limited funds. He says that he has seen an increase in bumper stickers in the scientist-heavy Washington DC area that read "Peer review isn't grant review. It's a lottery.". He agrees that reviewers seem to spend less time on each grant application and are now writing shorter comments — often in bullet points. He once valued feedback on rejected applications. "Now you get these trite little sentences."

When a proportion of a researcher's salary comes from grants, it is not surprising that people are seeing salary cuts. And in some cases, rising non-research costs, including outlays for health care, retirement, parking or mass transit, are upsetting to one-quarter of respondents, who say that they are adding to their job dissatisfaction. One researcher at George Washington University in Washington DC, who asked to remain anonymous, says that the amount he pays for parking has almost doubled in the decade he has been there. And his institution's health-insurance provider has raised its premiums yet decreased its coverage. Like some other US residents, he can benefit from his spouse's scheme, too, but not all researchers in the United States have this luxury. He says that these changes do not affect how "satisfied" he is with his job, but he knows that they trouble some colleagues.

DISPROPORTIONATE EFFECTS

Such costs bite especially deeply for early-career scientists, who tend to have smaller base compensation. Dominick Burton, a British postdoc at the Weizmann Institute of Science in Rehovot, Israel, needs to pay $800 a year for health insurance — a requirement for his employment. Burton says that the additional outlay (he would not have to pay anything in Britain) has pushed down his level of job satisfaction to satisfied rather than very satisfied.

A lucky 14% of respondents across all age-groups and career stages report that they are very satisfied with their job. Adil Mardinoglu is one of them. The Turkish native lives on a slim postdoc stipend and sometimes puts in 100-hour weeks at Chalmers University of Technology in Gothenburg, Sweden, but derives "positive energy" from his work on malnutrition in African children. "We are doing something good," he says.

That outlook may explain why many scientists report satisfaction yet bemoan funding and salary issues. "Everyone would always like more money for research," says "very satisfied" Thomas Merritt, a biochemist at Laurentian University in Sudbury, Canada. "But we're so lucky to get paid to do what we do, you can't spend the time whining about 'We need more, more, more.'" ■

Paul Smaglik *is assistant editor of* Nature *Careers. Additional reporting by* **Karen Kaplan**, **Shirana Kelly** *and* **Dan Penny**.

MENTORING

When guidance is important

The guidance that researchers receive about their work — whether from superiors or co-workers — contributes to their level of satisfaction. But in *Nature*'s latest salary survey, most respondents gave less than glowing reviews. Just one in four say that they are happy with the amount of guidance they have received in the past year, and half say that it has had little effect.

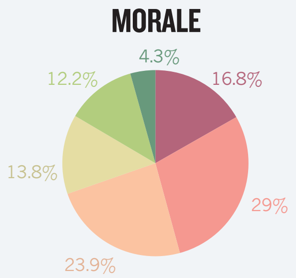
The responses seem to differ greatly by country. People in Japan gave the lowest ratings, with just 13% giving a thumbs-up. Conversely, one-third of respondents from the United States and Canada say that they are pleased with the level of guidance they have received. The difference could reflect the dissimilar cultures. Many US institutions have formal mentorship programmes and some federal grants require descriptions of the applicant's mentoring plans for junior scientists in their lab; in Japan, however, there are systemic issues that can hinder the proliferation of great mentors (see *Nature* **462,** 948; 2009).

Not many respondents think that they have sufficient opportunities for advancement, either. Fewer than one in three say that such opportunities had boosted their satisfaction in the past year, and two in five say that it has detracted from it. Just one in five participants from the United Kingdom, one in four from the United States and one in three from Japan — the nations with the most responses — say that they feel positive about career advancement. Across income levels, nearly half of those earning between US$40,000 and $69,999 — those most likely to be at the early stages of their careers — say that they are unhappy with advancement opportunities. **K.K.**

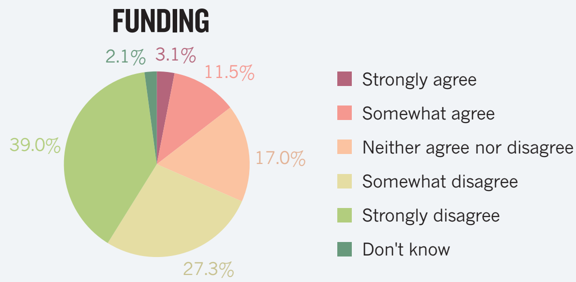
Money and morale

The 2014 *Nature* Careers salary survey collected 7,216 responses from researchers of every career stage around the world. Difficulty in securing grants correlates strongly with decreased lab morale, but most scientists are still happy with their jobs.

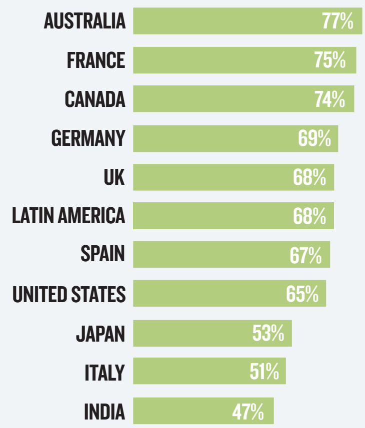
Q I feel that in the last year, the morale in my lab/department has decreased.



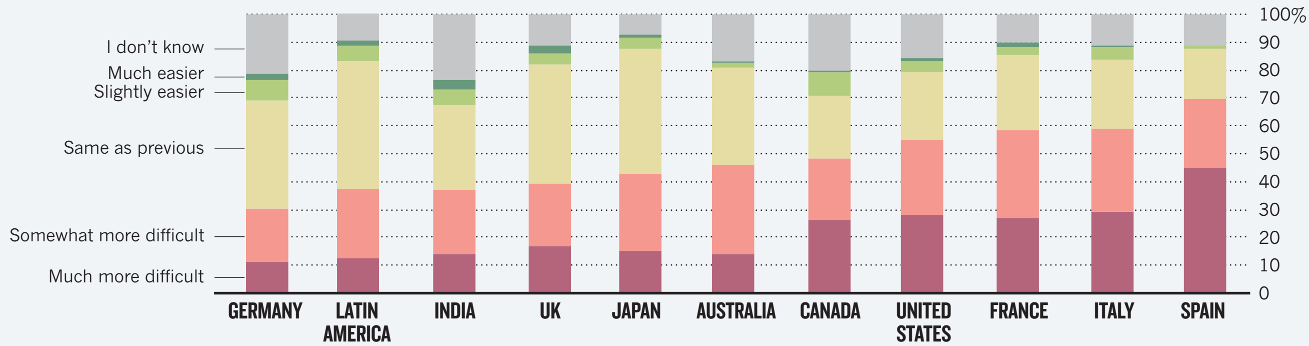
Q I feel that the science funding environment is improving.



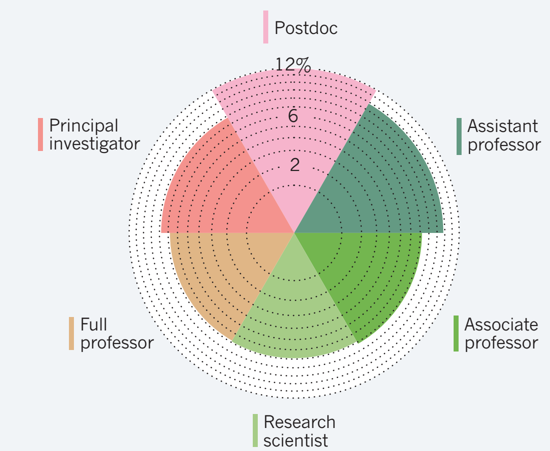
Q I am satisfied or very satisfied with my job.



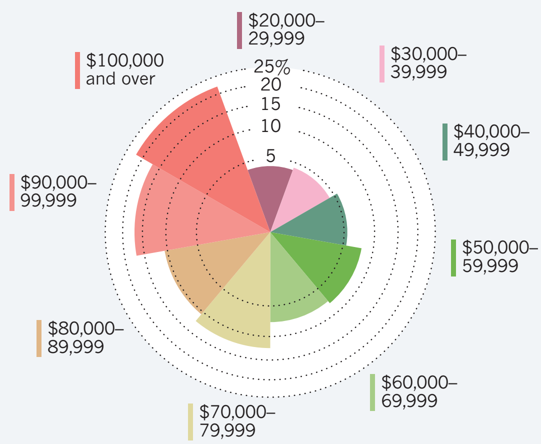
Q How did your efforts to secure grant funding in 2013 compare with your previous attempt?



Q In the past year, would you say your level of satisfaction has improved greatly?

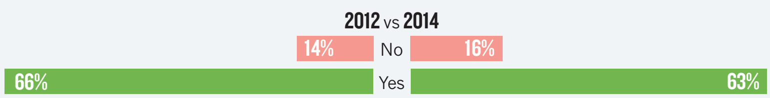


Q Would you say you are very satisfied with your salary?



4 OUT OF 10 RESEARCHERS ACROSS ALL NATIONS, INCOME AND POSITIONS SAY THAT THEIR SALARY NEGATIVELY AFFECTS THEIR JOB OUTLOOK

Q Overall, would you say you were satisfied with your current job?



THE TIGER WAITING ON THE SHORE

Days of remembrance.

BY PAUL CURRION

I am sitting in front of a man folded up like newspaper. His breath rustles in his bed sheets. His fluids are visible beneath his skin. The plastic pipe is more robust than the vein it feeds. I can hear so many things, including my own eyes blinking. I am looking at my son.

We have nothing to say to each other. At least, I have nothing to say to him, and he can't say anything at all. The last time I saw him, he was 14 years old, sitting on the other side of a perspex panel, and crying without control because I was about to be deep-sixed.

Now he is 94, and the only perspex panel between us is the passage of time. In 10 minutes, the prison guards will come and take me back to my bedroom. It's comfortable enough — more comfortable than this medical cell — but I won't be spending much time there.

This is the second day of my prison sentence. On the first day, I was woken by the guard and taken to my husband's wedding. I watched as he successfully moved on with his life without me. On the third day of my prison sentence, everybody I knew will be dead, and I will be released.

I accept my responsibility: involuntary manslaughter is still manslaughter. The family of the man I killed held a public vote, and the public spoke. The public is squeamish now — or at least, it was squeamish then, all those years ago — and so my punishment was humane.

Yet this is a hell. It's a hell that barely scratches the skin, true, a hell that puts me mercifully out of sight, but it is a hell that is merciless in its means. It is the afternoon of the second day of my prison sentence, and my punishment has barely begun.

I think for a moment about whether I should take my chances when the guards return; but death by cop is not an option any more, not with the next generation of non-lethal weapons, and who knows what advances they've made in the decades since I was sent away? Better not to risk it.

➲ NATURE.COM
Follow Futures:
@NatureFutures
go.nature.com/mtoodm

On the other hand, what more could they do to me? Let's imagine I managed to overpower the guards, escape the hospice, lose myself in the street somehow. I have no idea what the world looks like now, I have no idea what language they speak, or anything else considered common.

JACEY

No. I will go back to my bedroom, and go to sleep and wake up tomorrow, 100 years from now. I will pick up the package that all prisoners receive, and be released into a population of my peers. Criminals, like me. Exiles, like me. We will be housed, and fed, and cared for.

We won't be required to wear bracelets or chains. We will live our lives in some resource-constrained recreation of the society that put us to sleep, quarantined from whatever society we wake into. That future society may not want us, but I hope that they will be … humane.

When the result of the vote to decide my fate was announced, my husband — who began the process of successfully moving on at that very moment — tried to explain to my son that it wasn't so bad, that it could be worse, that I could have received the death penalty.

My son saw the truth that my husband could not see. Death was not for the squeamish, not then, but the alternative was far worse. That first night, I slept for a decade: the second night, 70 years; my third and final night may last 100 more; and I will know every second of that century.

I will be unawake, my body slowed by drugs, but my mind stretched over time as if it was on the rack. I will have 100 years to reflect on what I did, and what the consequences were, and my one memory will be of this: my own son, dying.

He shuffles himself into a different position, every card in his pack faded and creased. His mouth works like the mouth of a furious little animal that lives beneath the earth, thin and pale and cracked. He knows that I am here. He is trying to speak.

I lean forward so that I can hear him. These might be his last words. They will be his last words to me. "I found the family —" he starts to say, and then stops, and then, "— the man you killed." When he opens his eyes, the light of 80 years past shines again for me.

"It makes no difference," I tell him gently. "They think that this is punishment, but seeing you before you died — this is something I wanted." I reach for him — a breach of protocol — and he rubs his fingers against the back of my hand.

"They tried to reverse the decision," he tells me, and we laugh together. The very idea, of turning back time! Of putting the genie back in the bottle, of bringing back the dead! Time runs like a river for him, and crawls like a glacier for me — but still, it goes in one direction only.

Everybody knows that the past is a foreign country, but so is the future. I have been sent far away with no chance ever to return. Tomorrow I will wake, washed up on the shore of a brave new world, where I will stagger to my feet, and press on into the forest, to face whatever tigers await. ■

Paul Currion *is his own worst pseudonym. (More reliable information can be found at www.currion.net).*

nature OUTLOOK

LUNG CANCER

11 September 2014 / Vol 513 / Issue No 7517

Cover art: Russell Cobb

Lung cancer occupies a peculiar place in the public mind. It takes the lives of more people than any other cancer (page S2), yet, because the disease is so closely associated with the lifestyle choice of cigarette smoking, sympathy for its victims tends to be mixed with blame.

But neither lung cancer's inevitable end nor its attachment to cigarettes are accurate portrayals. There are embers of hope for new therapeutics. Some of the most promising developments come from therapies that turn the body's immune system against the disease (S10). On another front, drugs that target the genetics of particular tumours are emerging from the laboratory (S8). Screening technology can pick up tiny lung nodules when they are more easily treatable, although putting such screening into widespread use faces economic and institutional obstacles (S4 and S7).

The link between smoking and lung cancer has been firmly established for decades. And although most lung cancer can be attributed to direct inhalation of tobacco smoke, about one quarter of lung-cancer cases worldwide occur in people who have never smoked (S12) and who have arrived at their fate through some unlucky combination of genetics and environmental factors. Evidence is mounting that outdoor air pollution can cause lung cancer — findings that ought to spur action on reducing emissions, especially of particulates (S14). In Asia, lung cancer is alarmingly common in non-smoking women — apparently as a result of heavy use of indoor cooking stoves in unventilated homes (S16). In short, avoiding cigarettes, while still a wise health choice, is no guarantee against lung cancer.

This Outlook was produced with the support of Boehringer Ingelheim and Cancer Research UK. As always, *Nature* retains sole responsibility for all editorial content.

Herb Brody
Supplements Editor

CONTENTS

S2 EPIDEMIOLOGY
The dominant malignancy
Lung cancer by the numbers

S4 DIAGNOSIS
Early warning system
Screening for lung cancer could buy years of life for some patients

S7 PERSPECTIVE
The screening imperative
John K. Field is in favour of a widespread lung-cancer detection programme

S8 PERSONALIZED MEDICINE
Special treatment
The ups and downs of overcoming drug resistance using targeted medication

S10 IMMUNOTHERAPY
Chemical tricks
Can a new generation of antibodies outwit tumour cells?

S12 AETIOLOGY
Crucial clues
Analysing tumours in never-smokers is helping to reveal how the disease works

S14 ENVIRONMENT
Breathing trouble
Outdoor air pollution looks set to become a cancer-causing time bomb

S16 PUBLIC HEALTH
A burning issue
Why do so many non-smoking east Asian women develop lung cancer?

COLLECTION

S18 **Comprehensive molecular profiling of lung adenocarcinoma**
The Cancer Genome Atlas Research Network

S26 **Lung cancer chemoprevention: current status and future prospects**
Keith, R. L. & Miller, Y. E.

S36 **Rare variants of large effect in *BRCA2* and *CHEK2* affect risk of lung cancer**
Y. Wang et al.

S42 **Oncogenic and drug-sensitive *NTRK1* rearrangements in lung cancer**
A. Vaishnavi et al.

Nature Outlooks are sponsored supplements that aim to stimulate interest and debate around a subject of interest to the sponsor, while satisfying the editorial values of *Nature* and our readers' expectations. The boundaries of sponsor involvement are clearly delineated in the Nature Outlook Editorial guidelines available at go.nature.com/e4dwzw

CITING THE OUTLOOK
Cite as a supplement to *Nature*, for example, Nature **Vol. XXX**, No. XXXX Suppl., Sxx–Sxx (2014).

VISIT THE OUTLOOK ONLINE
The *Nature Outlook Lung Cancer* supplement can be found at http://www.nature.com/nature/outlook/lungcancer
It features all newly commissioned content as well as a selection of relevant previously published material.

All featured articles will be freely available for 6 months.

SUBSCRIPTIONS AND CUSTOMER SERVICES
For UK/Europe (excluding Japan): Nature Publishing Group, Subscriptions, Brunel Road, Basingstoke, Hants, RG21 6XS, UK. Tel: +44 (0) 1256 329242. Subscriptions and customer services for Americas – including Canada, Latin America and the Caribbean: Nature Publishing Group, 75 Varick St, 9th floor, New York, NY 10013-1917, USA. Tel: +1 866 363 7860 (US/Canada) or +1 212 726 9223 (outside US/Canada). Japan/China/Korea: Nature Publishing Group — Asia-Pacific, Chiyoda Building 5-6th Floor, 2-37 Ichigaya Tamachi, Shinjuku-ku, Tokyo, 162-0843, Japan. Tel: +81 3 3267 8751.

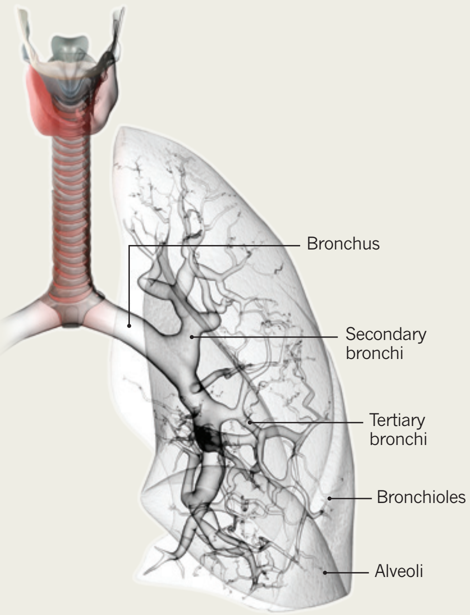
CUSTOMER SERVICES
Feedback@nature.com

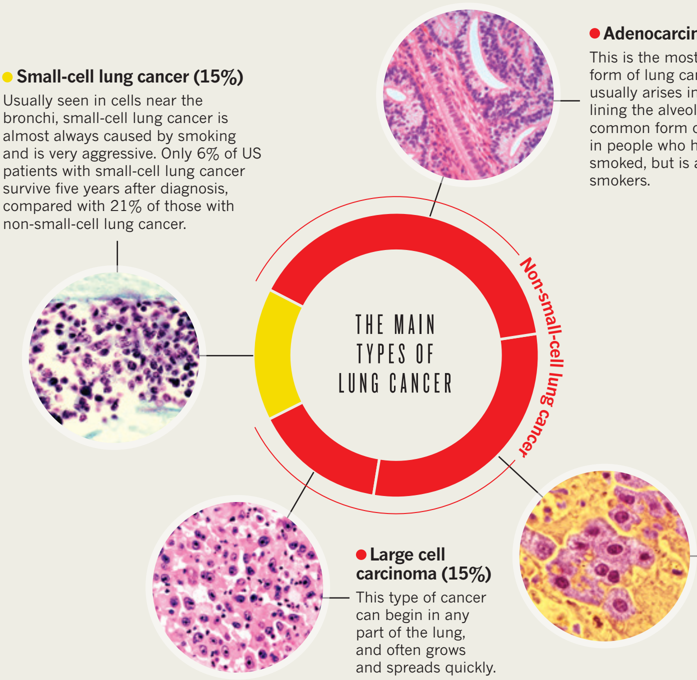
THE DOMINANT MALIGNANCY

*Lung cancer is the leading cause of cancer mortality. In some countries, incidence rates are dropping but survival rates for those with the disease remain low. By **Eric Bender**.*

MARKS OF A KILLER

Lung cancer is categorized by cell type into non-small-cell lung cancer — of which the three main subtypes are adenocarcinoma, squamous cell carcinoma and large cell carcinoma — and small-cell lung cancer. Treatment and prognosis differ depending on the type of lung cancer[1].





Small-cell lung cancer (15%)
Usually seen in cells near the bronchi, small-cell lung cancer is almost always caused by smoking and is very aggressive. Only 6% of US patients with small-cell lung cancer survive five years after diagnosis, compared with 21% of those with non-small-cell lung cancer.

Adenocarcinoma (40%)
This is the most prevalent form of lung cancer and usually arises in the cells lining the alveoli. It is a common form of lung cancer in people who have never smoked, but is also seen in smokers.

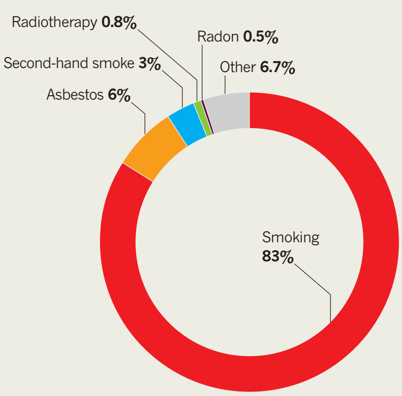
Squamous cell carcinoma (30%)
These tumours appear in the flat cells that line the inside of the airways, usually near the bronchi. This form of the disease is usually caused by smoking and is more common in men than women. The tumours tend to grow slowly.

Large cell carcinoma (15%)
This type of cancer can begin in any part of the lung, and often grows and spreads quickly.

PUBLIC-HEALTH CASE STUDIES

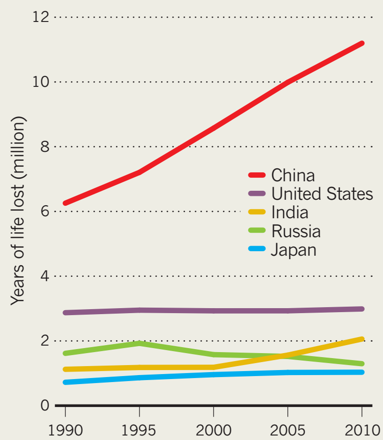
United Kingdom

Risk factors
Estimated causes of lung cancer in Britain, 2014[2]. Most cases of lung cancer are attributable to smoking, and so could be prevented.



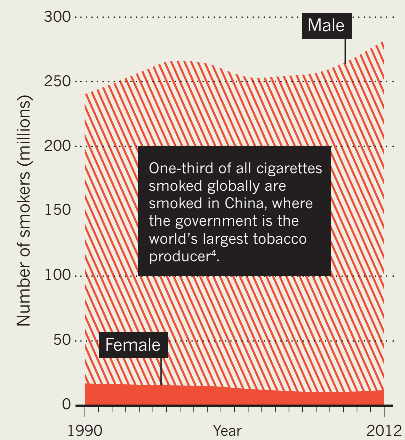
Worldwide

Lost years
The impact of the disease is climbing quickly in China, as a result of the rapidly increasing number of smokers (see 'Smoke rises', right)[3].



China

Smoke rises
The population of smokers, dominated by men, continues to rise in China, where tobacco kills an estimated 3,000 people a day[3].



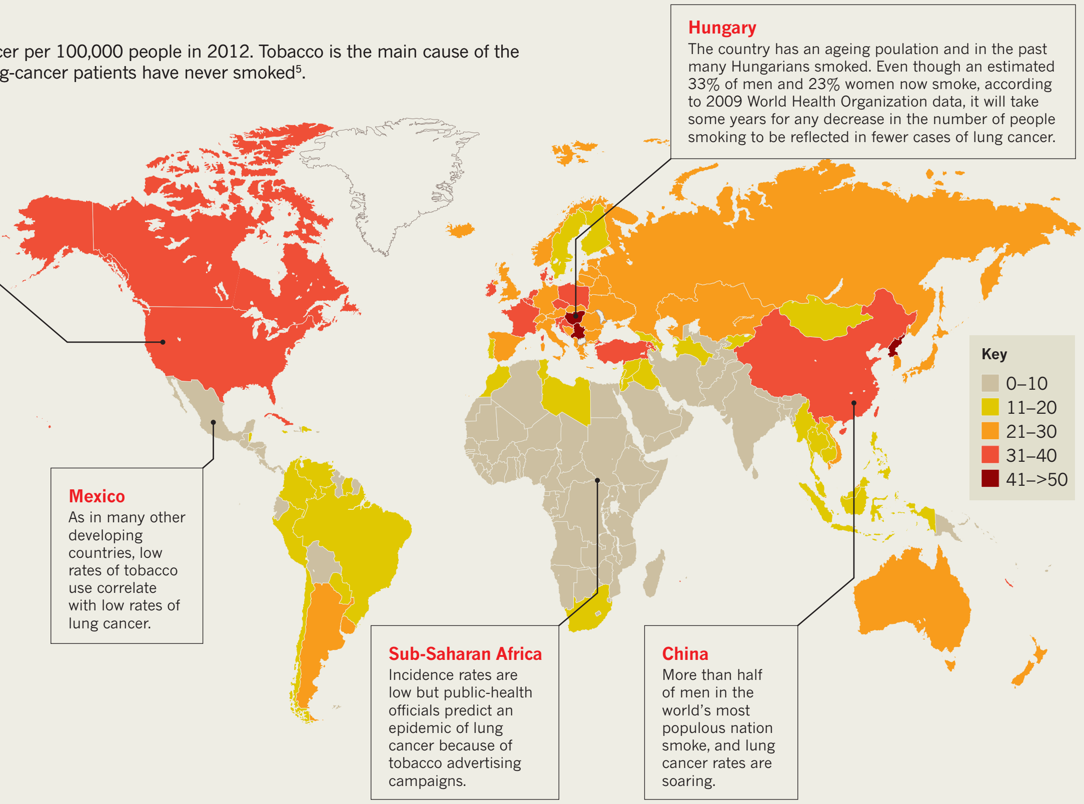
REFERENCES: 1. US National Cancer Institute, *Surveillance, Epidemiology, and End Results Program Cancer Statistics Review 1975–2011*; 2. Cancer Research UK; 3. Institute for Health Metrics and Evaluation, University of Washington; 4. International Agency for Research on Cancer, *World Cancer Report* (2014); 5. International Agency for Research on Cancer; 6. Belcher, E. *Tob. Control* **19,** 325–330 (2010); 7. Same as ref 3;

17.8%

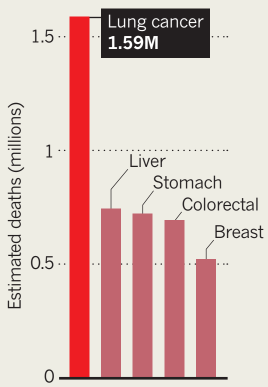
People living for five years after being diagnosed with lung cancer in the United States[1].

A WORLD OF RISK

Annual incidence risks of lung cancer per 100,000 people in 2012. Tobacco is the main cause of the disease, but about one-tenth of lung-cancer patients have never smoked[5].



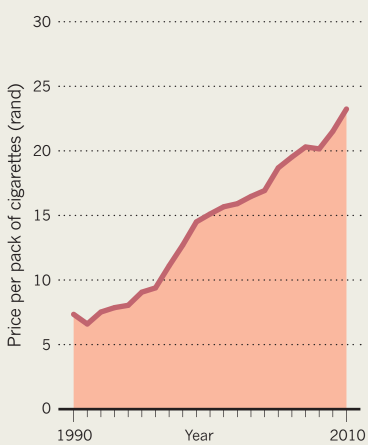
Estimated worldwide cancer mortality, 2012

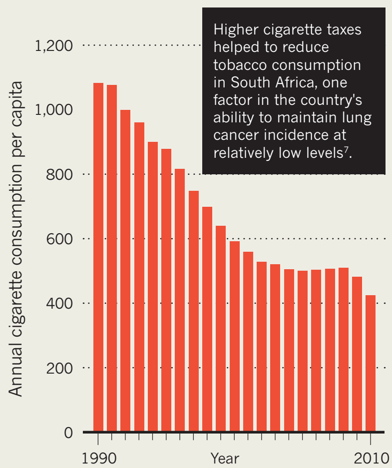


South Africa

Cigarette consumption and cost

Around the world, higher taxes on tobacco lower smoking rates. In South Africa[6], overall retail prices have risen substantially since 1991 and cigarette sales have plummeted.

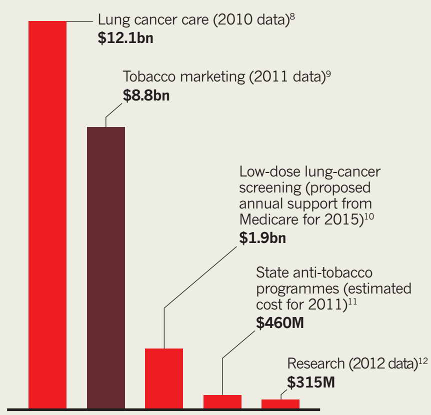




United States

Annual price of lung cancer

A debate continues over how much screening would cost, but proposed support from Medicare is a fraction of the spend on tobacco marketing.



8. US National Cancer Institute *Cancer Trends Progress Report 2011/2012*; 9. US Centers for Disease Control and Prevention; 10. Roth, J. A. *et al. J. Clin. Oncol.* 32 (Suppl.), 6501 (2014); 11. United States Centers for Disease Control and Prevention; 12. Same as ref 8.

SIEMENS HEALTHCARE

Computed tomography can detect tumours early, but is not yet widely used for lung-cancer screening.

DIAGNOSIS

Early warning system

The costs of lung-cancer screening overshadow the benefits of swift diagnosis — but ingenious technologies could help.

BY KATHERINE BOURZAC

At the age of 56, Gordon Green, a former smoker with two young children, was referred to a lung-cancer screening programme by his primary-care doctor, even though at the time he reported no health problems. A low-dose computed tomography (CT) scan that took less than a minute revealed a nodule in his lung that turned out to be a small, early-stage tumour. Doctors removed the growth and, two years later, Gordon is cancer free.

In patients such as Gordon, whose tumours are detected early, doctors see the potential that screening has to transform lung cancer from essentially a death sentence into a treatable disease. One of the reasons why lung cancer is so lethal is that diagnoses tend to be made after the cancer has advanced to late stages. Data collected by the US National Cancer Institute from 2004 to 2010 indicate that just 17% of people diagnosed with lung cancer are alive five years later. But it is not all bad news — some people who show no symptoms but whose cancer is detected in time have an 88% chance[1] of living another full decade, says radiation oncologist Andrea McKee, who runs the screening programme that detected Gordon's cancer at the Lahey Hospital and Medical Center in Burlington, Massachusetts.

Lung-cancer screening is not widely available anywhere in the world outside clinical trials and pilot programmes such as Lahey's, but that may be about to change. In a 2011 clinical trial in the United States[2], screening by low-dose CT reduced deaths from lung cancer by 20%. Based on these results, and on a recommendation by the US Preventive Services Task Force (an independent scientific body that advises the government), private insurers in the United States will have to start covering the costs of screening in January 2015 for current and former smokers aged 55–80 who are classed as high-risk: people with a history of smoking the equivalent of at least one pack of cigarettes a day for 30 years and, if former smokers, have quit smoking less than 15 years ago.

Yet there is uncertainty about who would foot the bill: Medicare, the national insurer for Americans aged 65 and older, is evaluating whether to pay for lung-cancer screening given the uncertainty about the procedure's cost-effectiveness. And CT screening comes with a high rate of false positives: Green had cancer, but 19 of 20 people in the same risk group — current and former smokers aged 55–74 — who were referred for further screening or biopsies did not test positive for cancer, and therefore underwent unnecessary and potentially risky surgical interventions.

Against this background, scientists and engineers are working on technologies to supplement, and perhaps eventually replace, CT scans. Too often, what seems to be a tumour is a harmless spot, so researchers are developing software to help to extract more accurate data from the images. Other research teams are evaluating biomarkers — biochemical or genetic indicators such as anti-cancer antibodies — in blood, sputum and even breath to ensure that healthy people are not sent for unnecessary biopsies.

After decades of poor outlooks for patients, the imminent availability of screening will change what it means to have lung cancer, says McKee. So far it's available only under the aegis of academic early adopters, but where it is in place, "I see a shift happening", she says.

SCREEN TEST

The research indicating that low-dose CT screening can lead to a 20% reduction in lung-cancer mortality in the United States was the outcome of a study[2] by the National Lung Screening Trial (NLST). The team compared screening (chest X-ray and low-dose CT) with non-screening in a nine-year, 53,000-patient randomized study that was completed in 2011. The trial administered three annual screening scans to half the participants, who were aged 55–74 when the trial started and were randomly selected to receive CT scans or X-rays. Participants were all classified as high risk.

Some scientists crave definitive information about lung-cancer screening. For example, Pierre Massion, a pulmonologist who runs the Nashville Lung Cancer Screening Trial at the Vanderbilt-Ingram Cancer Center in Tennessee, would like to know the effect of changing the screening interval from one to two years. But answers will have to wait: at a cost of about US$250 million, the large, complex trial needed to assess this is not likely to be done in the United States or anywhere else in the near future.

CT images are made by combining multiple X-ray images taken from different angles. Low-dose CT involves taking fewer images. The combined result is coarser than the images that are needed to diagnose blood clots, for example, but they are good enough to reveal lung nodules that warrant further investigation. Low-dose CT exposes patients to 1.5 millisieverts of radiation — about half the average person's annual radiation exposure, and about the same as a person would get on a single transatlantic flight. An annual low-dose CT scan is considered safe by the American College of Radiology, the organization that accredits radiology centres.

> *"Computers can pick up differences in these images that are too subtle for a radiologist."*

The NLST results led the US Preventive Services Task Force to recommend low-dose CT screening for adults with a history of smoking. Medicare does not have to follow the recommendations made by the task force, but under the Affordable Care Act, private insurers do. The American Medical Association, the American College of Radiology and the American Lung Association all support screening using low-dose CT. But Medicare is still undecided: in April 2014, the Medicare Evidence Development and Coverage Advisory Committee — a panel of doctors and other medical professionals — made a non-binding recommendation against screening after a series of short presentations by pulmonologists. Their individual reasons varied but many were concerned about the risks of false positives.

Some doctors think of Medicare coverage for lung-cancer screening as a matter of social equality. They note that most Americans in the age group in which lung cancer is most prevalent are on Medicare (according to the National Cancer Institute, the median age at diagnosis is 70). If screening is restricted to those who can pay out of pocket, it is not equitable, says Peter Bach, a lung-cancer risk specialist at the Memorial Sloan Kettering Cancer Center in New York City, who is a strong advocate for screening.

There is a lot of money at stake. Screening can detect cancer at a much earlier stage, when treatment is less expensive and more likely to save the patient's life. But early diagnosis is not always enough to offset the costs of imaging and diagnostics. "Screening is not going to save us money," says James Mulshine, a specialist in translational medicine at Rush Medical College in Chicago, Illinois, who served on the International Association for the Study of Lung Cancer's screening advisory committee.

The NLST has not yet published the results of its cost analyses, but other researchers have been trying to fill in the billion-dollar blanks. At the annual meeting of the American Society for Clinical Oncology in May 2014, researchers at the Fred Hutchinson Cancer Research Center in Seattle, Washington, presented preliminary estimates. Joshua Roth, a health economics and epidemiology researcher who led the study, cautions that their analysis considers only the price of the screening and not the gain to society of a person living a longer, more healthy life. Assuming that the incidence of lung cancer in the Medicare population is about the same as it was in the NLST group, the first five years of lung-cancer screening would cost the government an estimated \$9.3 billion, or about \$1.9 billion per year. (Mammograms cost Medicare an annual \$1.1 billion and prostate-cancer tests cost \$500 million.) Roth says that lung-cancer screening costs are likely to fall when more test centres open and screening becomes routine, as they did with mammograms.

STRINGENT THRESHOLD

One of the benefits of low-dose CT is that it is very sensitive — a distinct advantage when doctors are looking to detect tumours early on. However, the test also detects harmless nodules, inflammation, scars from past infections and other lesions that turn out to not be cancer. Of 100 nodules flagged for additional screening in the NLST, only 4 turned out to be tumours.

False positives place a tremendous psychological burden on a patient and put healthy people at risk of complications — and even death — following unnecessary biopsies, says preventive medicine specialist Jonathan Samet at the University of Southern California's Keck School of Medicine in Los Angeles. Older people, who tend to have more underlying health problems such as heart disease, are the most vulnerable. Nonetheless, he notes, an American Lung Association committee that he chaired published a report in April 2012 recommending Medicare screening even after factoring in these concerns.

One way to reduce the number of false positives is to make the standards for detecting a positive nodule more stringent. In the NLST, nodules with a diameter of 4 millimetres or larger were considered positive; all patients scoring positive were sent on for further diagnostic tests. Increasing the size threshold should help to reduce false positives without causing a dangerous delay in the detection of true cancers, according to a study published in 2013 by a multi-institutional research effort, the International Early Lung Cancer Action Program[3]. Its finding suggests that the threshold could be raised to 7 or 8 millimetres. Roth says that his group is using the NLST data to predict the lives saved and costs lowered if Medicare were to recommend a similar cut-off.

European lung-cancer specialists are also tackling the false-positive issue. Researchers running the 15,000-person lung-cancer screening trial (NELSON) based at Erasmus Medical Centre in Rotterdam, the Netherlands, factored the false-positive problem into the design of their 12-year trial[4], which is on target to wrap up late in 2015. If a person in the trial has a small lung nodule that doctors suspect is a tumour, a biopsy is not taken immediately. Instead, the patient returns for follow-up scans and a biopsy is performed only if the nodule has grown sufficiently in that time. NELSON's threshold for a positive CT is based not on the diameter of the nodule, as in the US studies, but on volume and growth rate: a biopsy is done if the nodule is larger than 500 cubic millimetres, or has doubled in volume in 400 days or less. Harry de Koning, a specialist in screening at Erasmus who runs NELSON, says that in the United States the screening recommendations are appropriate given the results of the NLST. But he contends that it would be advisable to wait to implement screening in Europe until the trial is finished, because incidence and risk differ between populations.

READING THE RESULTS

A successful screening programme relies on first-rate interpretation of results. Robert Gillies, chairman of cancer imaging at the Moffitt Cancer Center in Tampa, Florida, says that current interpretive methods, no matter how they classify the nodules, are not gleaning as much detailed information from the CT images as they could. Gillies hopes to cut CT false positives in half by using more sophisticated software. Radiologists typically factor in whether the nodule is calcified — usually a sign that it is benign — as well as its size, location and a few other features.

In an effort to improve the interpretation of the scans, Gillies has developed software called radiomics, which bases its analysis on 400 quantitative features from CT scans taken of people with lung and head-and-neck cancer. Radiomics factors in features such as shape and texture, which Gillies says may significantly improve the ability to discern whether a nodule is malignant. "Computers can pick up differences in these images that are too subtle for a human radiologist to see," Gillies says. His group has used the software to analyse existing data sets in which the outcomes are known. In retrospective studies run on the NLST database, for example, the software predicted which patients had cancer with 79% accuracy[5].

Gillies' group was awarded a \$1.6-million grant this year from the state of Florida to establish infrastructure for a screening programme, which began in July, that incorporates automated image analysis. It is important to expand screening programmes to as large a population as possible, says Gillies. However,

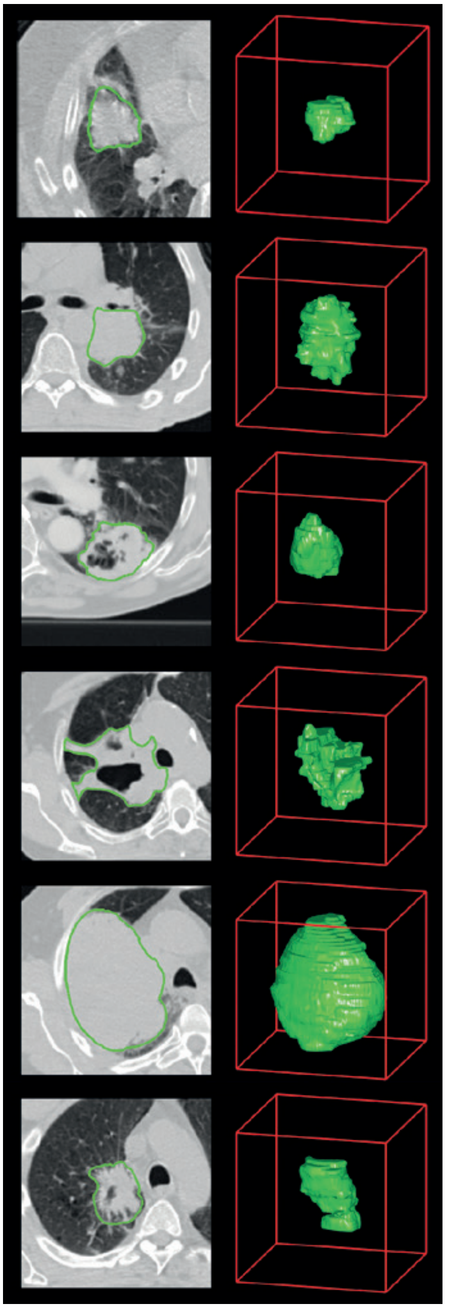
even though the tens of thousands of patients in the NLST provide his group with plenty of data to process, when it comes to making predictive models Gillies says that there is not actually a huge amount of information to work with. Because automated analyses get better only by being fed more data, Gillies wants states and countries to develop one database to share among researchers from across the globe. A good model to build on is a system run by the American College of Radiology called Lung-RADS. Although currently used as a quality assurance tool to standardize the reporting of screening results, Lung-RADS could one day integrate the sophisticated automated analyses that Gillies' group works with. Gillies is working on scaling up the Florida project as a testbed, but broader implementation of such a system will hinge on a favourable decision from Medicare.

BIOMARKER BONANZA

It might yet be possible to screen for lung cancer without using CT scans. Alternative approaches that could lower screening costs and increase patient safety include analysing blood and breath biomarkers, and detecting mutations in the nasal passages of cancer patients. Several smaller screening trials are under way to help to develop such screening technologies. For example, Avrum Spira, a pulmonologist and chief of computational biomedicine at the Boston University School of Medicine in Massachusetts, is looking at patterns of gene expression in cells taken from the airways, which can be sampled with brushes, a procedure that is safer than a biopsy of the lung. There are many changes in gene expression in the upper respiratory system that are associated with lung cancer[6] but their predictive value is as yet unproven. A clinical trial will determine the value of adding these markers to image-based patient workups.

Sam Hanash, a pathologist specializing in early cancer detection at the University of Texas MD Anderson Cancer Center in Houston, wants to speed up the process of validating biomarkers and getting the results to patients. Hanash's group has reviewed the literature for promising candidates and is embarking on a large validation study that will, he says, evaluate hundreds of biomarkers to see which ones are most effective in predicting cancer. To help push the screening techniques into practical use, Hanash says that his group is going to systematically test an array of biomarkers. "We'll do a bake-off, and figure out what is the best combination," he says.

The initial study will test markers in conjunction with CT screening. But ultimately Hanash hopes that biomarkers — which include proteins, antibodies and DNA — will replace CT screening. The study at the MD Anderson Cancer Center will start in the United States, but Hanash says the plan is to extend it to other countries, including China and Brazil, and enrol at least 10,000 people — current and former smokers who will be enrolled in the CT screening programme and have blood, sputum and other samples taken periodically. People who meet NLST criteria for screening would receive CT scans on enrolment and then again after the first and second years; then they would receive follow-up appointments for two additional years. At each visit, Hanash says, blood will be drawn and biomarkers observed. The reality check — seeing which patients develop lung cancer — will make clear the biomarker panel's false-positive and false-negative rates. It will cost about $100 million and take two years to screen enough people to do the validation, says Hanash. "This is going to be a huge undertaking."



Radiomics software extracts additional data from computed tomography scans to help doctors diagnose lung tumours (2D scans from six people, above, were used to build 3D images, in green).

As CT screening is introduced in academic centres, it is "changing the way we understand lung cancer", says McKee. One assumption has been that lung cancers are highly aggressive and must be removed once detected. As innovative screening methods enable earlier discovery of tumours, however, it is becoming clear that this may not always be the case. "In our study, one-third of lung cancers are indolent," says William Rom, director of pulmonary and critical-care medicine at the New York University School of Medicine. "When we remove them after five years, they're still stage 1." That means there is probably time to monitor people's health and perform follow-up imaging studies instead of sending them for biopsies or surgery. In the future, as cancers are spotted before they become aggressive, the type of treatment used will need to change — more surgery and radiotherapy followed by long-term care and monitoring, rather than late-stage diagnoses, expensive drug treatments and patient deaths.

> *"We'll screen, and we'll do it well — because the bar has been set really high."*

US doctors such as Bach and McKee are nervous but hopeful that Medicare will approve screening. "We will have to work hard to minimize harm if screening is implemented," says Bach. Doctors and insurers will need to make sure the scans are given only to the high-risk people who benefited from them in the NLST.

Meanwhile, delaying the start of lung-cancer screening means that more people may die from the disease when they could have been treated early and lived, as Green did, says McKee. The delay in the decision about lung-cancer screening from the Medicare advisory committee, and concerns about the cost, means that it is not being implemented rapidly, but with care, she says. "We'll screen, and we'll do it well — because the bar has been set really high." And if the US model for long-term screening saves lives, other countries may find it hard to resist following suit. ■

Katherine Bourzac *is a freelance science writer in San Francisco, California.*

1. The International Early Lung Cancer Action Program Investigators *N. Engl. J. Med.* **355,** 1763–1771 (2006).
2. The National Lung Screening Trial Research Team *N. Engl. J. Med.* **365,** 395–409 (2011).
3. Henschke, C. I. *et al. Ann. Intern. Med.* **158,** 246–252 (2013).
4. Horeweg, N. *et al. Eur. Respir. J.* **42,** 1659–1667 (2013).
5. Aerts, H. J. W. L. *et al. Nature Commun.* **5,** 4006 (2014).
6. Spira, A. *et al. Nature Med.* **13,** 361–366 (2007).

PERSPECTIVE

The screening imperative

Lung cancer kills more people than any other malignancy. Let's not delay in implementing a screening programme, says **John K. Field**.

Roughly every two minutes during 2012, someone in the European Union (EU) died of lung cancer. Those 268,000 lung-cancer fatalities represented more than one-fifth of all EU cancer deaths. The good news is that screening for lung cancer using low-dose computed tomography (CT) could reduce this enormous burden of mortality through early detection and treatment that improves survival[1]. Nearly 75% of lung-cancer patients present with late-stage disease, when effective treatment is unlikely to succeed. However, if the disease is treated at an early stage, more than 70% of patients survive another five years. Lung-cancer CT screening makes early detection possible, and so could add many years to many lives.

Unfortunately, there are major barriers obstructing the implementation of life-saving screening. The lung-cancer community has evidence for a mortality benefit from CT screening from a massive study in the United States: the National Lung Screening Trial (NLST). This randomized trial of more than 55,000 individuals — current and former smokers aged 55–74 — was stopped early when it became clear that low-dose CT screening resulted in a 20% decrease in lung-cancer mortality over screening with standard chest X-rays[2]. Based on these results, five clinical professional groups in the US support the implementation of CT screening, as does the US Preventive Services Task Force — although Medicare, the federal agency that insures Americans aged over 65, has not yet approved coverage.

EVERY YEAR WE **DELAY** COULD NEEDLESSLY **SACRIFICE** TENS OF THOUSANDS TO THE WORLD'S BIGGEST CANCER **KILLER**.

Despite the clear findings of the NLST, European health authorities have decided not to go ahead with lung-cancer screening. Instead, officials are awaiting the outcome of the NELSON trial[3] in the Netherlands and Belgium and the pooling of data from smaller EU trials, due in the next two years, which will provide European mortality and cost-effectiveness data[4].

This delay is a mistake. Now is the time to start planning to implement lung-cancer screening in Europe. The major stumbling block is uncertainty over screening's cost-effectiveness. In the US, lung-cancer screening is estimated to cost anywhere from US$19,000 to $160,000 per quality-adjusted life year (a standard method used to assess medical treatments by taking into account a person's quality of life after a medical intervention). But these figures are based on a health-care system that is very different from those that exist in Europe. Modelling in Britain, before the UK Lung Cancer Screening (UKLS) trial, provided an estimate of only £14,000 ($24,000) per quality-adjusted life year — a figure much more likely to be acceptable to a cost-conscious health-care system.

Clearing the cost hurdle is necessary but not sufficient for low-dose CT to be ready for widespread lung-cancer screening. Another issue relates to the criteria for interpreting the image produced by the scan. There are two schools of thought. One is to judge the nodule by its diameter, as measured by callipers on the radiograph. This is the approach used by the NLST. But diameter is not always accurate or consistent: nodules tend to be highly irregular. Thus a small nodule might show up as large if it is measured along its greatest dimension, creating a false-positive result, and vice versa for a large nodule measured along its shortest axis. That is why it's better to use the volume of a nodule to judge the risk it poses, which was what both the NELSON and UKLS trials did. This radiological approach has gained acceptance in Europe and is highly likely to reduce the number of false positives.

The next question to ask is: 'Who should be screened?' The US Preventive Services Task Force recommends that CT screening should be undertaken in past or present smokers aged 55–80 who meet the NLST entry criteria[5]. Evidence from the UKLS trial[6] — using the Liverpool Lung Project risk prediction model (people with a 5% risk of developing lung cancer in the next five years) however, shows that a screening programme will be more cost-effective if it is limited to the highest-risk segment of that population, which is those aged 60–75. Drawing a line like this will, of course, have life-and-death consequences; withholding screening from 55–59 year olds will result in a small number of lung cancers being missed. Such are the decisions that any preventative health programme must confront.

Likewise, there is no consensus on how often to screen. The largest evidence base is from the US trial, which involved annual scans. But modelling that uses the UKLS selection criteria and the NLST mortality data has shown that after an initial scan, the most cost-effective programme would involve not annual but biennial screening. According to this model, biennial scans would save 20% fewer lives than annual ones, but the predictions suggest that mortality benefit would still be substantial and cost effective[7].

The existence of unanswered questions about lung-cancer screening does not argue for inaction. The additional data that will flow out of the NELSON and pooled EU trials is necessary, but there is no need to wait before taking concrete steps towards planning to implement a widespread lung-cancer screening programme among the highest-risk populations. Every year we delay could needlessly sacrifice tens of thousands of lives to the world's biggest cancer killer. ■

John K. Field *is a clinical professor at the University of Liverpool Cancer Research Center, UK, and is the chief investigator for the UK Lung Cancer Screening trial.*
e-mail: J.K.Field@liv.ac.uk

1. Field, J. K. *et al. Lancet* **382,** 732–741 (2013).
2. National Lung Screening Trial Research Team *N. Engl. J. Med.* **365,** 395–409 (2011).
3. Van Klaveren, R. J. *et al. N. Engl. J. Med.* **361,** 2221–2229 (2009).
4. Field, J.K. *et al. Lancet Oncol.* **14,** e591–e600 (2013).
5. Humphrey, L. L. *et al. Ann. Intern. Med.* **159,** 411–420 (2013).
6. McRonald, F. E. *et al. Cancer Prev. Res. (Phila.)* **7,** 361–371 (2014).
7. Duffy, S. W. *et al. Br. J. Cancer* **110,** 1834–1840 (2014).

RUSSELL COBB

PERSONALIZED MEDICINE

Special treatment

Therapies targeted at the specific genetics of a patient's lung cancer have proved harder to realize than expected.

BY MICHAEL EISENSTEIN

Ramaswamy Govindan vividly remembers the first time he treated his patients with the cancer drug gefitinib. It was the start of the millennium, and the outlook for patients with metastatic non-small-cell lung cancer (NSCLC) was dire: less than 40% survived a year after diagnosis.

"The second patient I treated was about to go into hospice care," recalls Govindan, a medical oncologist at Washington University School of Medicine in St Louis, Missouri. "But she went on to live three years before dying of a heart attack."

Gefitinib was approved by the US Food and Drug Administration (FDA) in 2003. Marketed as Iressa by AstraZeneca, its arrival was a watershed moment in the treatment of NSCLC, the most common type of lung cancer. The drug blocks a protein called epidermal growth factor receptor (EGFR), which transmits signals that help to control the division and migration of cancer cells.

However, although some patients responded well to the treatment, many others did not. The same was true for another drug that targets EGFR: erlotinib (Tarceva), developed by Genentech and OSI Pharmaceuticals and approved by the FDA in 2004. The only apparent trend was that non-smokers were more likely than smokers to respond to erlotinib. "Back in the day, you would give Tarceva to somebody because they didn't smoke, but in the vast majority of those people it didn't help," says Mark Kris, a thoracic oncologist at Memorial Sloan Kettering Cancer Center in New York City.

In 2004, two research teams — one of which included Kris — discovered the secret[1,2]. Both gefitinib and erlotinib were selectively active against lung cancers with hyperactive, mutated versions of the *EGFR* gene, but ineffective against tumours in which the gene was not mutated. Mutated *EGFR* is predominantly found in a type of NSCLC called adenocarcinoma, which accounts for 40% of lung cancers and is the most common form of the disease in people who have never smoked.

The realization that specific genetic variants might help researchers to develop personalized lung-cancer treatments has launched a generation of targeted drugs that can deliver years of additional life to certain subgroups of patients. But some patients are still waiting to reap the medical benefits of the post-genomic era, and many doctors and clinical researchers fear that the low-hanging fruits of lung-cancer genetics may already have been picked.

FOLLOW THAT DRIVER

The cancer genome is a battered and scarred landscape of DNA-sequence changes as well as swapped, duplicated and deleted regions. The therapeutic focus is on the subset of these mutated genes — 'drivers' — that are essential for aggressive cell growth. The most useful drivers from a therapeutic perspective are oncogenes, which encode proteins that promote uncontrolled cell division and have the potential to convert a normally functioning cell into a cancer cell. Drugs that target mutant oncogenes might halt or reverse tumour growth.

One major lung-cancer oncogene is *EGFR*. Mutations to the *EGFR* oncogene are detected in more than 40% of adenocarcinomas. Three drugs are commercially available for *EGFR*-mutant cancers, and more are in trials. In 2007, researchers uncovered a second driver oncogene that is present in 5–7% of adenocarcinomas. Called *ALK*, this gene encodes a poorly understood signalling protein and occasionally undergoes a genomic rearrangement that leaves the resulting protein permanently turned on. In 2011, the FDA approved crizotinib (marketed by Pfizer as Xalkori) for NSCLC patients whose tumours exhibit such rearrangements. Phase III trial data presented by Pfizer at the 2014 annual meeting of the American Society for Clinical Oncology (ASCO) indicate that crizotinib can extend the life of patients whose tumours have mutations in *ALK*.

However, the benefits of these targeted drugs are only temporary — after about a year of remission, most tumours acquire resistance.

For example, more than half of the tumours treated with *EGFR* inhibitors acquire a mutation called T790M in the *EGFR* gene[3]. This blocks the drug without interfering with the mutant protein's signalling.

Tumours often contain genetically distinct cell populations, and many researchers believe that cancer recurrence may represent the evolutionary victory of an already-resistant minority. "Once we start to kill off cells that have the sensitizing mutation, the intrinsically resistant cells start to grow," says Tony Mok, a clinical oncologist at the Chinese University of Hong Kong.

Presentations at this year's ASCO meeting revealed promising clinical-trial data on drugs being developed by Clovis Oncology and AstraZeneca that inhibit the T790M mutant receptor. One molecule induced tumour shrinkage in almost two-thirds of patients.

Patients with crizotinib-resistant tumours also received hopeful news this year. Such resistance often arises in the absence of a detectable mutation, which suggests that other mechanisms increase *ALK* activity to overwhelm crizotinib's modest capacity for inhibition. In April 2014, the FDA moved with unprecedented speed to approve the drug ceritinib (marketed by Novartis as Zykadia) based purely on a phase I trial[4] showing a strong clinical response in resistant patients. Subsequent data suggest that ceritinib works equally well in both previously untreated and crizotinib-resistant patients.

Ceritinib is 5 to 20 times more potent than crizotinib as an *ALK* inhibitor, and it is also more selective, says Alice Shaw, an oncologist at Massachusetts General Hospital in Boston, whose team led the phase I trial. At least nine other *ALK* drugs are in development.

A FACE IN THE CROWD

Targeted treatments benefit only a minority of lung-cancer patients. For the rest, the hunt continues for drivers that might prove vulnerable to therapy. Most progress has been seen in people diagnosed with adenocarcinoma and who do not smoke, many of whom have cancers that have arisen through one primary driver mutation (see page S12). By contrast, the mutational load in a smoker's tumour can be overwhelming, making it a challenge to separate the signals of likely driver mutations from the noise generated by large numbers of 'passenger' mutations that make a minimal contribution to tumour growth.

But even targeting the genetic culprit in a single driver mutation can be tricky. Take the example of the oncogene *KRAS*, which encodes a signalling protein involved in cell proliferation. *KRAS* mutations appear in as many as one-quarter of adenocarcinomas, but attempts at targeted therapy have so far failed. A study reported at the 2014 ASCO meeting suggests that a subset of patients with *KRAS*-mutant NSCLC may benefit from a combination of drugs that target several proteins in the same biological pathway as *KRAS*. So far, only 10–15% of *KRAS*-mutant tumours respond to combination treatment, says Vassiliki Papadimitrakopoulou, a medical oncologist at the MD Anderson Cancer Center in Houston, Texas, who helped to coordinate the study. "We would like to see more than that."

For patients with non-adenocarcinoma lung cancers, targeted options are limited. Very few patients with squamous cell carcinoma (SCC) — the second most common form of lung cancer — have *EGFR* or *ALK* driver mutations. Most SCC tumours occur in smokers, and are plagued by the same extensive genomic mutation that is confounding efforts to apply targeted treatment to smokers' adenocarcinomas.

> ***"If you have lung cancer in 2014, the first thing we do is a genetic test for potential drivers."***

This may be about to change, thanks to the work of Govindan and his colleagues at the Cancer Genome Atlas (TCGA), which in 2012 published a detailed assessment of the SCC genomic landscape derived from tissue samples from 178 SCC tumours[5]. The results suggested a number of avenues for potential intervention. A mutation in the gene *CDKN2A*, for example, is found in 70% of SCC tumours and could be a target.

JOINT FORCES

The urgent need for progress in lung-cancer treatment has inspired Papadimitrakopoulou, who is collaborating with other US investigators on the Lung Cancer Master Protocol. Launched in June, this multi-arm, multi-institutional clinical trial will use sequencing to match SCC patients with targeted drug candidates. It will also accumulate a lot of cancer genomic data. "We will be characterizing the largest set of SCCs across the United States," says Papadimitrakopoulou.

Govindan and his colleagues are also working on large-scale genomic analysis. After a genomic survey of mutations in 230 adenocarcinoma tumours[6], published in July 2014, he and fellow TCGA coordinators Louis Staudt and Matthew Meyerson are working on plans to study a larger number of tumour samples in the hope of detecting additional targetable drivers.

The robust performance of drugs that target *ALK* and *EGFR* has made testing for mutations in these genes routine. But as the cost of sequencing plummets, some clinicians believe that it makes more sense to survey hundreds of cancer-related genes rather than just those two to provide a larger set of potential targets. Kris is among the evangelists for extensive clinical sequencing. "If you have lung cancer in 2014, the first thing we do is a biopsy that includes a comprehensive genetic test for all potential drivers," he says. Companies are also providing the tools to do this. Foundation Medicine, a company in Cambridge, Massachusetts, co-founded by TCGA scientists, generates oncology diagnostic reports for clinicians based on sequencing data from 236 cancer-associated genes. The company expects to do 25,000 tests in 2014, up from 9,000 in 2013. In June, the Memorial Sloan Kettering Cancer Center forged a partnership with Quest Diagnostics of Madison, New Jersey, to broaden clinician access to the centre's in-house genetic test, which also surveys numerous oncogenes in parallel.

Genetic analyses could help to identify patients with mutations that are rare in lung cancer but are common in other tumour types. For example, a subset of adenocarcinoma patients with mutations affecting the *RET* gene might benefit from cabozantinib, a drug that targets this alteration in thyroid cancer[7]. And with much of the pharmaceutical industry's oncology efforts focused on developing targeted drugs, data from sequencing the genes of lung-cancer patients can also help to direct those patients to clinical trials. To assess the impact of sequencing on lung-cancer care, Kris and other scientists — who formed a group called the Lung Cancer Mutation Consortium — sequenced as many as 10 known oncogenes in more than 1,000 patients. Kris reports that 28% of the people tested were matched to clinical trials they might not otherwise have known about[8].

As with *KRAS*, many oncogenes are informative scientifically but are not medically useful, leading some researchers to question the short-term benefits of routine, large-scale tumour sequencing in patients — a practice Mok says is unlikely to improve lung-cancer care significantly until the next *EGFR* comes along. Still, he believes that genetic analysis must be embedded into the diagnostic process so that drugs can be matched to a patient as quickly as possible — he holds out hope that new drivers will soon join *ALK* and *EGFR*.

As would everyone struggling to find new weapons against this lethal disease. With such resources at hand, more doctors might look forward to experiencing the sweet satisfaction Govindan encountered on providing his patient with just the treatment she needed to buy years of additional life. ■

Michael Eisenstein *is a freelance science writer in Philadelphia, Pennsylvania.*

1. Lynch, T. J. *et al. N. Engl. J. Med.* **350,** 2129–2139 (2004).
2. Pao, W. *et al. Proc. Natl Acad. Sci. USA* **101,** 13306–13311 (2004).
3. Pao, W. *et al. PLoS Med.* **2,** 73 (2005).
4. Shaw, A. T. *et al. N. Engl. J. Med.* **370,** 1189–1197 (2014).
5. The Cancer Genome Atlas Research Network *Nature* **489,** 519–525 (2012).
6. The Cancer Genome Atlas Research Network *Nature* **511,** 543–550 (2014).
7. Drilon, A. *et al. Cancer Discov.* **3,** 630–635 (2013).
8. Kris, M. G. *et al. J. Am. Med. Assoc.* **311,** 1998–2006 (2014).

BENOÎT HELLER/TRANSGENE

The cancer vaccine TG4010, shown here being manufactured by Transgene, extended lung cancer survival time in a phase II trial.

IMMUNOTHERAPY

Chemical tricks

Lung cancer uses cunning mechanisms to evade the immune system. Can new antibody therapies outwit the disease?

BY BIANCA NOGRADY

The immune system has evolved over millions of years to protect the human body against microbes, pathogens and parasites. Which makes it all the more puzzling to immunologists as to why, when it comes to helping the body defend itself against cancer, immunotherapy treatments designed to enhance the immune system have so far failed to make even the slightest dent in halting the spread of the disease.

So when medical oncologist Naiyer Rizvi became involved with the phase I trial of a tumour antibody a few years ago, he was prepared for failure. In fact, there was a certain glum expectation in the lung-cancer community that this trial would go the way of so many other attempts to fight cancer by enlisting the body's own immune system.

One of the first trial patients Rizvi saw at Memorial Sloan Kettering Cancer Centre in New York City had a large adrenal tumour that was causing him so much pain he was rushed to hospital for emergency treatment soon after he got his first dose of the trial treatment, the immunotherapeutic agent nivolumab that was under development by Bristol-Myers Squibb based in New York City. In May 2014, Rizvi saw the same patient again. It was one year since completion of a two-year course of therapy with nivolumab and the man's tumours were still shrinking. "When you've got these dramatic unexpected responses," Rizvi says, "you kind of rethink the direction of your career."

He's not the only one feeling this way: a wave of optimism is sweeping through the lung-cancer field. Data from trials of different immunotherapies raise the promise of new agents with response rates and survival advantages that outweigh anything else on offer, adding months and even years to life expectancy.

CHECKPOINT CHECKMATE

There is a long, sad history of immune system approaches to cancer therapy going awry. Early attempts to develop drugs that would help immune cells fight tumours failed dismally in clinical trials. Vaccines would generate the desired immune response, and there would be high levels of immune cells primed to attack the malignancy. Yet for reasons that researchers could not understand, there appeared to be no effect on tumours: they weren't shrinking.

For Lieping Chen, an immunologist at Yale University in New Haven, Connecticut, it was the frustration of having so little to offer his oncology patients that drove him into the field of cancer immunology. That was more than two decades ago, when the first steps were being taken to discover how molecules on the surface of tumour cells might affect the body's immune response. Chen set out to find specific molecules that would stimulate an immune response against tumours or inhibit whatever was blocking that response.

Chen's first big success came in the late 1990s, when his team discovered and cloned the gene coding for a protein that keeps the body's immune response in check — a fundamental function because if the immune system enters overdrive, the result would be chronic and damaging inflammation[1]. Not long after Chen's discovery, a separate group figured out the mechanism by which this protein kept the immune system from over-responding. The key was in its interaction with T cells (white blood cells that are a key component of the immune system). Specifically, Chen's protein interfered with a receptor on the surface of T cells called PD-1. By binding to this receptor, the protein triggered the death of the T cell, suppressing immune activity[2].

Further research showed that the T cells are involved in their own demise (see 'Immune assistance'). When T cells arrive at their target, they release a molecule called interferon-γ to boost their cell-killing ability. The tumour takes advantage of this mechanism — interferon-γ causes tumours to go into overdrive in their production of Chen's protein, now known as PD-L1. And when the PD-L1 protein binds to the PD-1 receptor on T cells it makes the T cells commit suicide.

The finding shed light on one of the main mechanisms that allowed tumours to neutralize T cells. "You can have a very good response in the blood and in the lymphatic organs, but they shut down in the tumour cell," says Chen, speaking about the immune system.

The goal, then, is to disrupt this pathway and therefore unleash the immune system to attack the tumour. A class of drugs called checkpoint inhibitors do this by one of two means. The first approach is to introduce a molecule that binds to the T cell's PD-1 receptor and therefore prevent the tumour protein PD-L1 from doing so. As a result, T cells can resume their normal function and destroy the tumour.

Nivolumab, which Rizvi used so effectively at Memorial Sloan Kettering, is an antibody that does just that. Other researchers have also seen its impressive results. Julie Brahmer, an oncologist at the Johns Hopkins Sidney Kimmel Comprehensive Cancer Center in Baltimore, Maryland, has seen nivolumab move from phase I through to phase III trials. She says that at least two of her patients, both treated with nivolumab for two years, are still alive a year and a half after their treatment finished. Her patients had advanced, metastatic lung cancer and had already undergone treatment with chemotherapy and radiation, so the odds were stacked against them. "Both of these patients should no longer be around if you look at the statistics around lung cancer," says Brahmer.

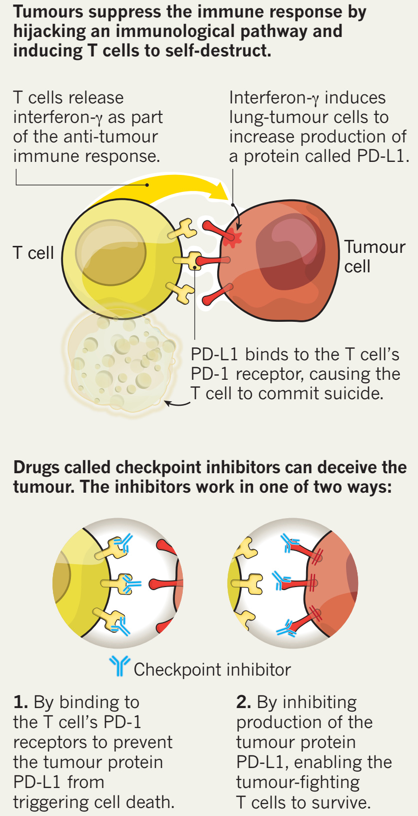
The other way to interfere with this pathway is not to bother with binding to the receptor, but instead block the tumour's PD-L1 protein directly. That's what medical oncologist Jean-Charles Soria is doing. A phase I trial designed to test the safety and clinical activity of an antibody that blocks PD-L1 found that around 25% of patients with non-small-cell lung cancer (most lung cancers are of this type) responded to the drug. Soria, who heads drug development at the Gustave Roussy Cancer Center in Paris (and who reported the results of the ongoing study at the 2014 American Society of Clinical Oncology meeting) notes that this response rate is better than the 3% rate generally seen in patients who are receiving their third course of chemotherapy after earlier treatments have failed.

Although the response rates are similar for drugs that bind to the T cell's PD-1 receptor and those that block the tumour's PD-L1 protein, there may be a slight safety advantage in targeting PD-L1. The phase I nivolumab study reported a 3% incidence of drug-related pneumonitis[3] — inflammation of the lung tissue — but this side effect has so far been less severe or absent with the PD-L1 inhibitors.

Both experimental therapies seem to benefit smokers more than never-smokers. Soria reported the results of a phase I trial of the PD-L1 inhibitor at the 2013 European Cancer Congress (see go.nature.com/b3z1wr); the study indicated that 26% of smokers responded to the drug, but only 10% of never-smokers

IMMUNE ASSISTANCE

New drugs help to outwit lung tumours.



SOURCE: D. M. PARDOLL *NATURE REV. CANCER* **12**, 252–264 (2012).

responded. Researchers speculate that this is probably due to the greater number of mutations present in smokers' tumours, an abundance that would probably present the newly awakened immune response with a far greater array of tumour antigens to respond to.

Despite the positive results from tests of both varieties of checkpoint inhibitors, Brahmer says that researchers are keeping their optimism in check. "Long-term disease control is probably the most realistic expectation of these antibodies," she says, likening the outlook to that of treatment for a chronic condition such as hypertension.

GLIMMER OF HOPE

The main purpose of checkpoint inhibitors is to undo the local blockade of the immune response and allow the immune system to resume normal function and attack the tumour, but they might also breathe life into vaccine therapies against lung cancer.

Philippe Archinard is chief executive of drug-development company Transgene, of Illkirch-Graffenstaden, France, which is preparing for phase III trials of its therapeutic lung-cancer vaccine, TG4010. He believes the two approaches are complementary. "Boosting the immune system is one way to address the issue," he says. But to achieve the greatest benefits, "you would need to increase the immune response and create it when it's not present".

The TG4010 vaccine uses a viral vector (a tool used to transport genetic material into a cell) that has been engineered to manufacture a variant of a tumour glycoprotein called MUC1. When MUC1 is injected into the body, the aim is to stimulate the immune system to respond to and attack cells that contain it. The MUC1 antigen is widely expressed on the surface of non-cancerous epithelial cells. But many tumours overexpress an abnormal version of it, which is why it is such a potent target for immunotherapy. A phase II trial of TG4010, used in conjunction with standard chemotherapy in patients with advanced non-small-cell lung cancer, showed that patients given the vaccine as well as chemotherapy survived a median of 17.1 months compared to 11.3 months for patients on chemotherapy alone.

TG4010 is not the only vaccine targeting MUC1. Pharmaceutical firm Merck Serono, based in Darmstadt, Germany, under a license agreement with Oncothyreon of Seattle, Washington, is developing a vaccine called tecemotide. In a 2013 phase III trial, tecemotide's survival benefits did not quite reach statistical significance[4]. However, an analysis of a subgroup of patients showed that those who were undergoing both chemotherapy and radiation treatment concurrently survived for around ten months longer than patients given the placebo vaccine. The lead researcher on the tecemotide trial, Charles Butts, says that as more data accrue from the trial, they point to a continued survival advantage — there's an almost 10% improvement in three-year survival compared with placebo, says Butts, an oncologist at the University of Alberta, Edmonton, and at the Cross Cancer Institute in the same city.

After 20 years of treating lung cancer, and 20 years of dashed hopes in lung-cancer immunotherapy, Butts says there is finally a glimmer of hope in both vaccines and non-vaccine immunotherapies, such as the PD-1 and PD-L1 inhibitors. "The fact that these checkpoint inhibitors are showing such responses — and durable responses — is quite amazing," he says.

If the past decades of failure have taught the industry anything, it is that early trial success rarely leads to a breakthrough drug. Phase III trial data from the checkpoint inhibitors and the vaccines are eagerly awaited — and researchers are waiting anxiously to see if these treatments can deliver the results that lung-cancer patients and their doctors are so desperately hoping for. ■

Bianca Nogrady *is a freelance science writer in Sydney, Australia.*

1. Dong, H. *et al. Nature Med.* **5,** 1365–1369 (1999).
2. Freeman, G. J. *et al. J. Exp. Med.* **192,** 1027–1034 (2000).
3. Topalian, S. L. *et al. N. Engl. J. Med.* **366,** 2443–2454 (2012).
4. Butts, C. *et al. Lancet Oncology* **15,** 59–68 (2014).

RUSSELL COBB

AETIOLOGY

Crucial clues

Studies in never-smokers have revealed key lung-cancer mutations — but the cause of the disease is still a mystery.

BY SARAH DEWEERDT

The lung-cancer patients that thoracic oncologist Sébastien Couraud remembers most are those who have never smoked cigarettes. He recalls one woman who tried for years to get her husband to stop his heavy habit, but in the end it was her, not him, who developed lung cancer — perhaps from breathing second-hand smoke. Another patient, the wife of a smoker, developed lung cancer long after her husband died of the disease. Couraud also remembers a group of colleagues who had been exposed to the same workplace carcinogen and who attended chemotherapy treatments together — until one day one of them didn't. "It's these patients you keep in your mind," says Couraud, who works at Hospices Civils de Lyon in France.

About one quarter of lung-cancer cases worldwide occur in people who have smoked fewer than 100 cigarettes in their life. In Europe and the United States, people who have never smoked account for 10–15% of lung cancers. In southeast Asia, half of all the women who develop lung cancer have never smoked. In fact, if lung cancer in never-smokers were considered a distinct disease, it would be the seventh leading cancer killer worldwide[1] (see 'Killing without smoke).

It makes sense to consider lung cancer in never-smokers separately. "It is almost like a different disease," says Joan Schiller, a lung-cancer specialist at the University of Texas Southwestern Medical Center in Dallas. Lung cancer in people who have never smoked is almost always a subtype of non-small-cell lung cancer called adenocarcinoma. By contrast, smokers get not only adenocarcinoma but also squamous cell carcinoma and small-cell lung cancer. Tumours in never-smokers tend to be less aggressive than in smokers, although they are frequently diagnosed at a more advanced stage because never-smokers, and their doctors, regard lung cancer as an exceedingly unlikely prospect and so often miss the early signs.

Tumours in never-smokers also tend to carry a distinctive set of genetic changes called driver mutations that are involved in turning cells malignant. Classifying patients according to their history of smoking has helped to understand lung cancer's gene mutations over the past decade, but researchers have found that this is not the best strategy for treating individual patients. That is because the most effective treatment often depends on the molecular characteristics of the tumour, not the characteristics of the patient. "Smoking status is sort of a surrogate for that, but it's an imperfect surrogate," says thoracic oncologist Charles Rudin at Memorial Sloan Kettering Cancer Center in New York City. So the task now is not only to continue to work out the patterns and consequences of tumour mutations, but also to delve into some of the mysterious aspects of lung cancer in never-smokers — especially the genetic and environmental causes and how to mitigate them.

GENETIC VARIATIONS

Studying the mechanisms of lung cancer is easier in never-smokers because they have not been exposed to the onslaught of DNA-altering chemicals in cigarette smoke. This has helped researchers to sort out which changes in a lung-cancer cell are driver mutations and which are passenger mutations — those that are simply along for the ride. "The lung cancers that occur in never-smokers are genetically simpler," says Rudin. "They have fewer mutations, but they may have the key mutations that are really important drivers."

The first clues that studying lung cancer in never-smokers might be particularly helpful in understanding the mechanisms of the disease emerged in the early 2000s. Clinical trials analysing a class of cancer medication called small-molecule tyrosine kinase inhibitors, which targets a family of proteins that

are mutated in many types of cancer, showed that never-smokers, individuals with adenocarcinoma, women and people with east Asian ancestry were more likely to respond well to the drugs than people with a history of smoking[2].

In 2004, three independent groups published studies that uncovered the molecular basis behind these observations. This class of tyrosine kinase inhibitors is effective against lung cancers that carry mutations in the epidermal growth factor receptor (*EGFR*) gene[3-5]. These mutations are more common in lung cancers that occur in the groups that responded well to the drugs in clinical trials. *EGFR* mutations are seen in 28% of never-smokers with lung cancer in the United States and in 68% of Asian people. By contrast, such mutations occur in only 5% of current smokers and in 11% of former smokers with lung cancer in the United States.

Since then, researchers have identified additional lung-cancer driver mutations that are more common in never-smokers than smokers[1]. "Many of the discovery efforts have been focused on never-smokers as a way of finding these driver mutations," says Geoffrey Oxnard, a thoracic oncologist at Dana-Farber Cancer Institute in Boston, Massachusetts. Researchers have identified therapies that target some of these tumour mutations, and the search is on for others.

FATAL RESISTANCE

The relationship between the types of lung-cancer mutations and whether someone smokes are not absolute. For example, although *EGFR* mutations are more common in never-smokers, one-third of lung cancers with *EGFR* mutations occur in smokers — therefore, knowledge of driver mutations and corresponding treatments gleaned from studies of never-smokers may benefit smokers with the disease. Testing for mutations in genes such as *EGFR* is gaining popularity as a tool for lung-cancer management in smokers and never-smokers.

Half to three-quarters of lung-cancer patients who have never smoked carry at least one mutation that will respond to targeted therapies such as tyrosine kinase inhibitors. This might seem encouraging news for never-smokers with lung cancer — but only to a point. "Their cancer is more treatable than cancer in smokers and they live longer as a result of having these targetable mutations, but we're not curing them," says Barbara Gitlitz, a lung-cancer specialist at the University of Southern California in Los Angeles. "It's still an extremely deadly disease." In part, this is because of the lower lung scrutiny that never-smokers get. "We're diagnosing these people at stages where they're not curable," Gitlitz explains.

But there is more negative news: the targeted therapies that benefit many never-smokers with lung cancer eventually stop working because the tumours develop drug resistance. Tackling drug resistance, suggests Rudin, will require better versions of targeted therapies — or better ways to use them (see page S8).

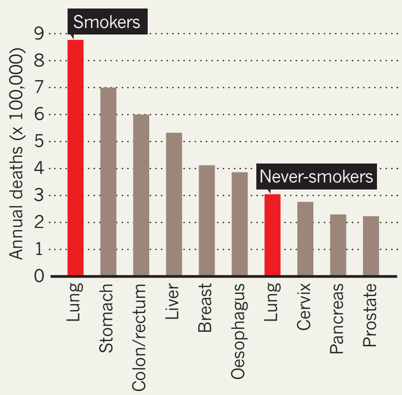
TROUBLESOME RISKS

Perhaps an even bigger mystery is what causes lung cancer in never-smokers, and how risk factors produce different driver mutations in lung tumours. "Lung cancer in never-smokers is a very interesting tool to focus on risk factors for lung cancer other than smoking," explains Couraud, who is working on a comprehensive study of tumour mutations among 384 never-smokers in France who have lung cancer.

Some risk factors are well known — breathing in second-hand cigarette smoke, for example, which is responsible for 20–50% of lung-cancer deaths in never-smokers in the United States. Studies have shown[1] that the more second-hand smoke a person is exposed to, the less likely he or she is to have *EGFR*-mutant lung cancer — in other words, breathing in a lot of second-hand smoke is likely to cause the same form of lung cancer as that seen in smokers. Curiously, however, data from the French cohort of never-smokers does not show this pattern — in fact, Couraud reports, those data show no relationship between second-hand smoke exposure and any driver mutation.

KILLING WITHOUT SMOKE

If considered as a separate disease, lung cancer in people who have never smoked would rank seventh in global cancer mortality.



SOURCE: REF. 2.

And tobacco smoke is not the whole story. In east Asia, never-smokers who develop lung cancer are disproportionately women, in part because of exposure to coal smoke in unventilated homes (see page S16). And in 2013, the International Agency for Research on Cancer confirmed outdoor air pollution as carcinogenic (see page S14). As the number of people smoking cigarettes continues to decline throughout the world, risk factors for lung cancer will change. "Lung cancer is not going to entirely go away because we convince people to stop smoking," Oxnard says.

Before cigarette smoking became widespread, lung cancer was rare, leading to just 0.7% of cancer deaths in the United States in 1914, versus an estimated 27% in 2014. Respiratory cancers — a category that includes not only lung cancer but also mesothelioma — are the most common cancers acquired as a result of exposures on the job to carcinogens such as asbestos or silica, for example.

"Lung cancer is not going to entirely go away because we convince people to stop smoking."

Connecting what is known about lung-cancer risk factors to individual patients remains difficult. "We don't have a clear understanding of why the majority of never-smokers develop lung cancer," Rudin says.

Some lung-cancer risk probably also comes from inherited genetic factors. Until five years ago, most studies investigating familial lung cancer have focused on families who smoked. As a result, there has been no good way to distinguish whether it is exposure to second-hand smoke or genes that have caused lung cancer. Researchers are just beginning to puzzle out the inherited factors that increase lung-cancer risk in the absence of exposure to tobacco smoke. A few studies have identified individuals with an inherited mutation in *EGFR*. This mutation, working through a mechanism that is not yet understood, seems to produce resistance to targeted therapies and also increase susceptibility to developing lung cancer[6,7].

The population of people who have never smoked but have lung cancer has become a model for studying other subgroups of people with the disease. Oxnard and Gitlitz, for example, are co-leading a study of genomic changes in patients who were diagnosed with lung cancer before the age of 40. Lung cancer is rare in this age group, and researchers say that studying this population may help to uncover additional driver mutations and therapeutic approaches — just as studies of never-smokers have done. "We as clinicians have the responsibility to keep our eyes open for such clinical outliers," Oxnard says, "because they may provide unique insights on a more deep biological level." ■

Sarah Deweerdt *is a freelance science writer in Seattle, Washington.*

1. Couraud, S. *et al. Eur. J. Cancer* **48,** 1299–1311 (2012).
2. Sun, S., Schiller, J. H. & Gazdar, A. F. *Nature Rev. Cancer* **7,** 778–790 (2007).
3. Lynch, T. J. *et al. N. Engl. J. Med.* 350, 2129–2139 (2004).
4. Paez, J. G. *et al. Science* 304, 1497–1500 (2004).
5. Pao, W. *et al. Proc. Natl. Acad. Sci. USA* 101, 13306–13311 (2004).
6. Oxnard, G. R. *et al. J. Thorac. Oncol.* **7,** 1049–1052 (2012).
7. Gazdar, A. *et al. J. Thorac Oncol.* **9,** 456–463 (2014).

PAUL SOUDERS/CORBIS

Air pollution is known to be linked with lung cancer, but it is uncertain how many people are in danger.

ENVIRONMENT

Breathing trouble

Large-scale studies are confirming suspicions that air pollution significantly increases the risk of lung cancer.

BY TRACI WATSON

There are many things people can do to lower their risk of developing lung cancer. They can choose not to smoke, they can stay away from smoky places and they can avoid breathing dirty air. Air pollution is often particularly high in cities, which are home to more than half the global population. With every breath, people living in polluted areas may inhale the same carcinogens as a smoker, including tiny contaminants called particulates.

Less than a decade ago, some researchers were doubtful that outdoor air pollution at levels common in the West could lead to significant health risks. But now a growing number of epidemiological studies is establishing a strong link between polluted air and lung cancer[1,2]. The International Agency for Research on Cancer (IARC) branded[3] outdoor air pollution — contamination from transportation, power generation, industrial and agricultural emissions, and heating and cooking — as carcinogenic in late 2013. At the same time, it designated particulate matter — microscopic droplets that are made from dust, the byproducts of power plants and other chemical components — as carcinogenic, after reviewing numerous studies linking it to lung cancer (see 'Something in the air').

With the link between air pollution and lung cancer now solid, researchers are embarking on studies to put numbers to the worldwide impact. Initial results indicate a much higher share of lung cancer linked to air pollution than previously thought — one analysis[4] suggests that in some countries, one form of particulate pollution was a factor in 10–20% of lung-cancer deaths. "Some estimates are larger than ours, some smaller," says epidemiologist Aaron Cohen of the Health Effects Institute in Boston, Massachusetts, who is one of the study's authors. "But all of them are large enough to be of concern for public health."

THE MURKY METROPOLIS

The epidemiological research needed to show an association between lung cancer and air pollution in humans is a monumental challenge. Most cases of lung cancer are associated with smoking, which makes it difficult for researchers to disentangle tobacco use from other potential causes. Lung cancer also takes time to develop, so investigators must follow a cohort of people for decades before a sufficient number of non-smokers will be diagnosed with the disease.

And air pollution in general is tricky to study. It is a complicated cocktail of chemicals that waft from sources from factories to lawnmowers. That complexity has frustrated researchers trying to discover what makes air pollution so carcinogenic. A number of studies show that fine particulate matter — particles less than 2.5 micrometres in diameter — is a key player. Placing blame on air pollution is also tricky because air quality differs between countries, cities and even neighbourhoods. Many early studies in the United States relied on skimpy networks of air-quality monitors and could not capture that variability. "It's a very hard area to study," says environmental health scientist Michael Jerrett of the University of California, Berkeley. "I tell my graduate students they should never study cancer in the environment until they've gotten tenure" and are no longer under such intense pressure to publish rapidly.

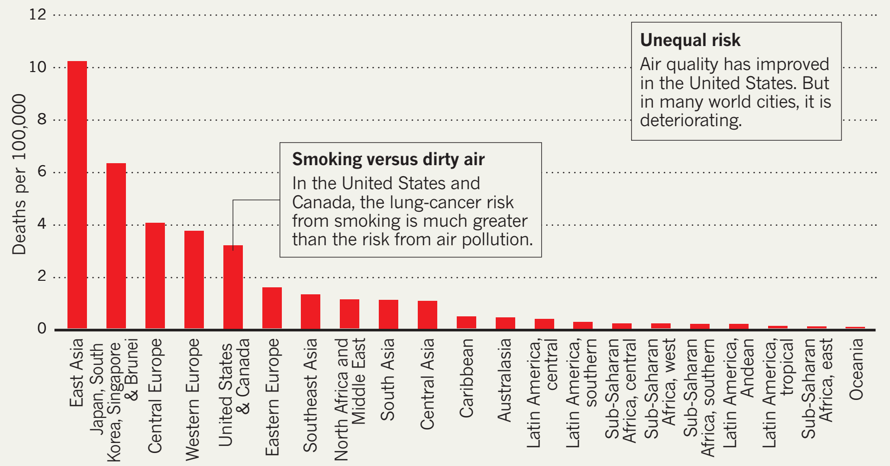
But researchers are learning to navigate some of the obstacles. An innovative technique called land-use regression uses monitors to measure air quality at dozens or even hundreds of locations. The data are then used to build an air-quality model that takes factors such as road length and topography into account to predict pollution levels at the address of every person in an epidemiological study — a vast improvement over the old days when researchers assumed everyone in a city was exposed to the same concentrations

of airborne contaminants. Land-use regression and related techniques "have definitely improved the evidence base a lot", Jerrett says.

These methods were deployed on a grand scale in a 2013 study across nine European countries[1]. The authors estimated exposure to pollution at the addresses of more than 300,000 people, including smokers and non-smokers. They found that a rise by 10 micrograms per cubic millimetre in fine particulate levels — roughly the difference between average pollution levels in a cleaner European city (Oslo) and a dirtier one (Athens) — increases the risk of lung cancer by 40%. The study, one of the largest epidemiological analyses of the link, confirmed previous results in US populations[5]. The results carried substantial weight with the IARC working group that evaluated the evidence linking air pollution with lung cancer, says working-group member Francine Laden, an epidemiologist at the Harvard School of Public Health in Boston, Massachusetts.

SOMETHING IN THE AIR

Global lung-cancer death rates in 2010 attributable to particulate matter, microscopic airborne droplets or particles that can be traced to sources including power-plant chimneys and dusty fields. The link between lung cancer and fine particulate matter — up to 2.5 micrometres in diameter — is especially strong.



SOURCE: AARON COHEN, INST. HEALTH METRICS AND EVALUATION, UNIV WASHINGTON.

NON-SMOKERS BEWARE

Just over a decade ago, Jonathan Samet, an epidemiologist at the University of Southern California in Los Angeles and the chair of the IARC working group, co-authored a review calling the epidemiological evidence for a link between air pollution and lung cancer "equivocal". But that uncertainty is dissipating rapidly. Scientists have conducted "studies that are really quite different kinds of designs, quite different types of populations, different exposure assessments, and we're still seeing these elevated risks", says Michael Brauer, an environmental health scientist at the University of British Columbia in Vancouver, Canada, who was also a member of the IARC working group. "You can't say this is an erroneous result that's driven by this design feature in this type of study."

To reassure themselves that air pollution is a risk factor for lung cancer, researchers needed data that ruled out exposure to cigarette smoke. Exactly such a study was published in 2011. The study examined nearly 189,000 people in the United States who had never smoked and took into account their exposure to second-hand smoke. Between 1982 and 2008, 1,100 of those people died of lung cancer, translating to an increased lung cancer-death risk of up to 27% for every 10 μg/m^3 rise in fine particulates[2].

Michael Thun, a co-author of the study and an epidemiologist at the American Cancer Society who is now retired and based in Atlanta, Georgia, confesses that before he collaborated on the study, he had doubts that low levels of air pollution in the United States were sufficient to cause lung cancer. But, he says, "the results in never-smokers substantially strengthened my confidence that the relationship was causal", and he is now convinced.

Thun is not the only person to change his mind. Samet now says that the epidemiology shows an association between air pollution and lung cancer, adding that results from studies of large populations are bolstered by laboratory findings and by research on people who breathe in high levels of air pollution at work. The link is now widely accepted by researchers in the field, Brauer says.

NO LOWER LIMIT

Researchers have yet to find a safe level of air pollution. The European study found a higher risk of lung cancer even at fine-particulate levels lower than the European Union's limit of 25 μg/m^3, and the authors found no evidence of a threshold — that is, no level of air pollution below which there is no increased risk of developing the disease. The analysis also found a linear relationship between incidence of lung cancer and air pollution levels, at least at concentrations common in developed countries.

Now that a link between air pollution and lung cancer has been established, researchers are building sophisticated models to determine the number of people around the world who develop lung cancer from exposure to it. However, because most research has been focused in areas with low levels of pollution, data investigating the risk posed by outdoor particulate levels of more than 30-40 μg/m^3 are scant. A team working on one study[5] sought to estimate the danger of breathing very high levels of particulates by using inhaled tobacco smoke as a proxy. Like the European study, their analysis showed a nearly linear relationship between lung-cancer risk and exposure to fine particulates, with no evidence of a threshold.

Another study[6] combined data indicating the risk posed by smoking and inhaling second-hand smoke as a way to assign risk numbers to countries, such as China, that have very high particulate levels. The results indicate that in many countries, fine particulates played a part in 10–20% of lung-cancer deaths in 2010. Based in part on the methods used in the paper, the Institute for Health Metrics and Evaluation, a research centre at the University of Washington in Seattle, found that in 2010, outdoor particulate pollution contributed at least in part to 223,000 lung-cancer deaths worldwide[7].

"The results in never-smokers strengthened my confidence that the relationship was causal."

The estimates published so far come with big uncertainties. No one really knows the size of the risk at the extremely high levels of pollution seen in cities such as Beijing, because most results to date have come from studies in developed countries. To produce firmer numbers, researchers need more data from countries with much more pollution.

It took a long time for scientists to establish the link between air pollution and lung cancer, but it may take a lot longer for that knowledge to save large numbers of lives. ■

Traci Watson *is a freelance science writer in Washington DC.*

1. Raaschou-Nielsen, O. *et al. Lancet Oncol.* **14,** 813–822 (2013).
2. Turner, M. C. *et al. Am. J. Resp. Crit. Care Med.* **184,** 1374–1381 (2011).
3. Loomis, D. *et al. Lancet Oncol.* **14,** 1262–1263 (2013).
4. Pope, C. A. III. *et al. Environ. Health Perspect.* **119,** 1616–1621 (2011).
5. Pope, C. A. III. *et al. J. Am. Med. Assoc.* **287,** 1132–1141 (2002).
6. Burnett, R. T. *et al. Environ. Health Perspect.* **122,** 397–403 (2014).
7. Inst. Health Metrics & Evaluation, Global Burden of Disease Cause Patterns 2010, available at: go.nature.com/brc4nw.

Carcinogenic emissions from burning smoky coal in poorly ventilated homes present grave health risks.

PUBLIC HEALTH

A burning issue

An unusually high number of women from east Asia develop lung cancer. Few smoke, but that's only part of the mystery.

BY NIDHI SUBBARAMAN

Lung cancer snuck into the lives of Angela Tan (not her real name) and those she loved. When she died her family and friends were stunned. The Chinese schoolteacher who lived in Singapore was in her early 40s and had never smoked a cigarette in her life.

Tan's death seemed to challenge the prevailing assumption that lung cancer is a smoker's disease. In fact, she had plenty of company: non-smoking east Asian women have been diagnosed with the disease in high numbers. "There is a striking disconnect between smoking rates and lung cancer incidence, particularly in Chinese women," says Adeline Seow, a cancer epidemiologist from the Saw Swee Hock School of Public Health at the National University of Singapore.

Seow has another connection to this story: she is Tan's niece. She was a medical student when her aunt died, and it left an indelible mark on her. "At the time, there was no explanation that we were aware of," she recalls. So she set out to find that explanation and answer the question: what makes east Asian women particularly susceptible to lung cancer?

Cancer researchers have called this phenomenon a natural experiment. It's a chance to study the aetiology of lung cancer that arises independently of its most notorious causative agent: tobacco smoking, says Dean Hosgood, who studies cancer and population health at the Albert Einstein College of Medicine in New York City. "As tobacco smoking is decreasing, other factors are going to become a larger proportion of lung cancer cases," he says. That shift makes investigating this group all the more valuable — and researchers are racing to identify the pathways involved in the disease, tease out the molecular signature of its mutations and flag genetic variants that make some people unusually prone to developing lung cancer.

INHERITANCE MATTERS

Lung cancer kills about 1.6 million people around the world every year, and is responsible for the largest number of cancer deaths worldwide. In the past decade, researchers have recognized that as many as 25% of lung cancers occur in people who have never smoked. Studying the issue closer, they've come to recognize that lung cancer in smokers has a distinct genetic signature from the disease seen in smokers (see page S12).

The incidence of lung cancer among non-smoking women also varies with geographical location, epidemiological studies have shown. It has been estimated that non-smoking women in east Asia are four times as likely to develop lung cancer than women in Europe or Africa (see 'Women's risk').

Seow's own research has focused mainly on lung cancer in urban Singaporeans. She has found that since the 1960s, cancer rates among Chinese immigrants (including smokers and non-smokers) are higher than among Malays and Indians, the other major ethnic groups in Singapore.

But the reported increase in incidence of lung cancer is not exclusive to the Chinese. In the past few years, several studies[1] have shown that people with Japanese and Korean ancestry also carry heritable genetic variants that put them at risk. To further study this phenomenon and consolidate data from across the region, cancer epidemiologists Qing Lan and Nathan Rothman at the United States National Cancer Institute (NCI) established the Female Lung Cancer Consortium in Asia in 2009.

Just five years later, the consortium has borne fruit, with implications that extend well beyond the tiny island-nation of Singapore. In 2012, Lan, Rothman, Seow and a dozen other researchers published the results of a genetic study involving more than 14,000 women — 6,600 of whom had never smoked but had lung cancer, and 7,500 controls — from 6 countries in east Asia. In an analysis known as a genome-wide association study — in which researchers look for genetic variations that occur more frequently in people with a disease — the team compared genome markers found in people with lung cancer to genome markers in the cancer-free control group. They found[2] three new DNA sites associated with the disease — one on chromosome 10 and two on chromosome 6 — and confirmed the relationship between three other variants previously flagged on chromosomes 3, 5 and 17.

Such a study would have been very difficult without the concentration in east Asia of non-smokers with lung cancer. "To get the numbers to look for genetic causes, conducting studies in Asia has a big advantage," says

Neil Caporaso, a specialist in the genetics and epidemiology of lung cancer at NCI and a co-author of the study.

The evidence also allowed the researchers to rule out the influence of another gene variant, on chromosome 15, that had been a suspected player in non-smokers' lung cancer. Its absence in this large cohort confirmed other data that suggested that people with the variant at a particular location on this chromosome, known as 15.25q, are more likely to develop cancer if they smoke tobacco.

It is also clear that the genes do not act alone. Variation at two particular regions on chromosomes 10 and 6 are associated with a 30% and 15% increase in the risk of developing lung cancer, respectively. (By contrast, smoking increases risk by 2,500%.) It is likely that genes work alongside an environmental factor, perhaps something in the air, says Hosgood, who was a postdoctoral researcher with Lan and Rothman, and a co-author on the study.

FANNING THE FLAMES

For many years, Xuanwei, a farming town in western China, had one of the highest rates of lung-cancer incidence anywhere in the world. As early as the 1970s, researchers noticed that women in the community were developing lung cancer at equal or higher rates as men, despite being less likely to smoke. In the mid-1980s, a handful of teams investigated the cause, and quickly found a suspect: coal.

Judy Mumford, a researcher with the US Environmental Protection Agency, was among the first investigators on the ground. Mumford and her colleagues explained that although women in Xuanwei rarely smoked tobacco, at home they cooked food over open fire pits that burned smoky coal[3]. Families also used coal to heat their houses and closed the shutters on their windows to keep the cold out, which trapped smoke and particulate matter inside. Chimneys were uncommon and many homes were poorly ventilated.

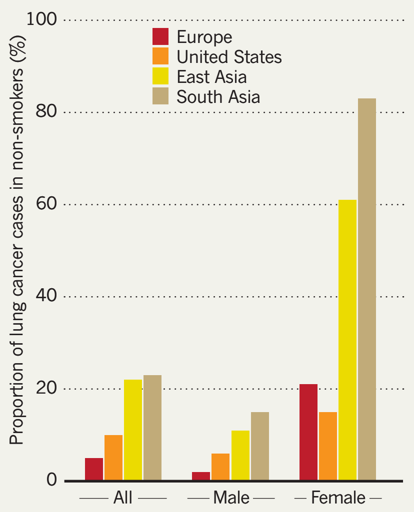
"What component of coal causes lung cancer? That's an important area that needs to be studied."

Ten years before Mumford's 1987 paper was published, Chinese authorities suspected that the mortality rate for women in Xuanwei from lung cancer was higher than anywhere else in the country. In fact, the death rate was 8 times higher than the national average and 17 times higher than the rest of the province[3]. But change was taking root. In 1976, alert to the unusually high death rates in Xuanwei, the Chinese government offered 10 yuan (US$5) to families to spend on building chimneys.

Lan was a graduate student when she first travelled to Xuanwei in the 1990s. Looking for a clear link between coal and cancer, Lan investigated the effect of the change in stove ventilation. She interviewed farmers who had lived in the area since 1976, and studied hospital records for the incidence of lung cancer. Her results were dramatic: in families that stopped cooking over an open fire indoors and heating their homes with smoky coal, lung cancer rates decreased by 41% in men and 46% in women[4].

WOMEN'S RISK

In Asia, non-smokers with lung cancer are mainly female. The discrepancy is explained by the fact that cooking indoors with coal is common.



SOURCE: REF. 6.

"That was a key study that really showed pretty convincingly that coal was related to lung cancer," says Rothman. In a follow-up study, Hosgood and colleagues showed that when families switched to a portable stove that could be used outside, men reduced their risk of dying from lung cancer by 39% and women by 59%[5]. Subsequent research on mice showed that indoor smoky coal emissions can be 1,000 times more carcinogenic than cigarette smoke[6].

In southern and eastern China, where coal use is less frequent, a different story emerged. Studies of home life in places such as Shanghai, Hong Kong and Taiwan made the case that cooking oils were to blame for higher lung cancer rates in those areas. Oils, including rapeseed, heated to high temperatures in woks for stir-frying food emit volatile carcinogens that can be inhaled.

In 2006, leaning on the overwhelming evidence presented in studies such as those published by Lan and Hosgood, the World Health Organization's International Agency for Research on Cancer formally addressed the carcinogenicity of coal and cooking oils. Nineteen scientists met in Lyon, France, and agreed that burning coal in households was carcinogenic to humans and that emissions from frying at high temperatures were probably carcinogenic to humans.

Researchers now know that repeated exposure to coal smoke in poorly ventilated areas can double the likelihood of developing lung cancer. But there is still more to be learned. "What component of coal causes lung cancer? That's an important area that needs to be studied," says Lan. She has early evidence that immunoregulatory genes involved in inflammatory pathways may play a part, but says larger studies are needed to establish this link.

A DIFFERENT DISEASE

Some clinical oncologists are making the case that lung cancer in never-smokers is a distinct disease and are starting to recognize its calling card. Never-smokers' lung cancers seem to be missing some of the genetic characteristics found in smokers' tumours. Mutations of the tumour suppressor gene *TP53* are abundant in cancers of all kinds, and frequently appear in the lung cancers of smokers. Yet, never-smokers rarely carry such mutations in their own cancer tissue.

On the flip side, researchers have found that 58% of lung cancers in non-smokers carry a specific mutation in the epidermal growth factor receptor gene, *EGFR*, compared with only 13% of smokers who have the same mutation. That the *EGFR* mutation is seen more frequently in lung tumour samples from east Asians than other populations — and more often in women than men — suggests that an associated mechanism will shed light on what makes east Asian women particularly vulnerable to the disease.

Continuing research is leading to a better understanding of lung cancer. Lan and other teams who have studied Xuanwei have established the carcinogenicity of coal smoke, spurring action to minimize the use of coal cookstoves in unventilated houses. They have also found that environmental factors are only part of the story. The abundant incidence of non-smokers' lung cancer among east Asian women has given researchers a rare opportunity to tease out the genetic variants that have a role in the disease. It is increasingly clear, says Seow, that the disease behaves differently in east Asian women.

For her own part, Seow has dual motivations for studying this group. One is from a public health perspective: understanding the dense web of risk factors will offer opportunities for lung cancer prevention. And her aunt's death all those years ago gave "a face to the disease", personalizing her efforts of the past two decades. Through her work, she hopes other Angela Tans can be saved. ■

Nidhi Subbaraman *is a freelance science writer in Somerville, Massachusetts.*

1. Daiki, M. *et al. Nature. Genet.* **42,** 893–896 (2010).
2. Lan, Q. *et al. Nature. Genet.* **44,** 1330–1335 (2012).
3. Mumford, J. L. *et al. Science* **235,** 217–220 (1987).
4. Lan, Q. *et al. J. Natl. Cancer Inst.* **94,** 826–835 (2002).
5. Hosgood, H.D. III *et al. Br. J. Cancer* **99,** 1934–1939 (2008).
6. Sun, S., Schiller, J. H. & Gazdar, A. F. *Nature Rev. Cancer* **7,** 778–790 (2007).

OPEN
doi:10.1038/nature13385

Comprehensive molecular profiling of lung adenocarcinoma

The Cancer Genome Atlas Research Network*

Adenocarcinoma of the lung is the leading cause of cancer death worldwide. Here we report molecular profiling of 230 resected lung adenocarcinomas using messenger RNA, microRNA and DNA sequencing integrated with copy number, methylation and proteomic analyses. High rates of somatic mutation were seen (mean 8.9 mutations per megabase). Eighteen genes were statistically significantly mutated, including *RIT1* activating mutations and newly described loss-of-function *MGA* mutations which are mutually exclusive with focal *MYC* amplification. *EGFR* mutations were more frequent in female patients, whereas mutations in *RBM10* were more common in males. Aberrations in *NF1*, *MET*, *ERBB2* and *RIT1* occurred in 13% of cases and were enriched in samples otherwise lacking an activated oncogene, suggesting a driver role for these events in certain tumours. DNA and mRNA sequence from the same tumour highlighted splicing alterations driven by somatic genomic changes, including exon 14 skipping in *MET* mRNA in 4% of cases. MAPK and PI(3)K pathway activity, when measured at the protein level, was explained by known mutations in only a fraction of cases, suggesting additional, unexplained mechanisms of pathway activation. These data establish a foundation for classification and further investigations of lung adenocarcinoma molecular pathogenesis.

Lung cancer is the most common cause of global cancer-related mortality, leading to over a million deaths each year and adenocarcinoma is its most common histological type. Smoking is the major cause of lung adenocarcinoma but, as smoking rates decrease, proportionally more cases occur in never-smokers (defined as less than 100 cigarettes in a lifetime). Recently, molecularly targeted therapies have dramatically improved treatment for patients whose tumours harbour somatically activated oncogenes such as mutant *EGFR*[1] or translocated *ALK, RET,* or *ROS1* (refs 2–4). Mutant *BRAF* and *ERBB2* (ref. 5) are also investigational targets. However, most lung adenocarcinomas either lack an identifiable driver oncogene, or harbour mutations in *KRAS* and are therefore still treated with conventional chemotherapy. Tumour suppressor gene abnormalities, such as those in *TP53* (ref. 6), *STK11* (ref. 7), *CDKN2A*[8], *KEAP1* (ref. 9), and *SMARCA4* (ref. 10) are also common but are not currently clinically actionable. Finally, lung adenocarcinoma shows high rates of somatic mutation and genomic rearrangement, challenging identification of all but the most frequent driver gene alterations because of a large burden of passenger events per tumour genome[11–13]. Our efforts focused on comprehensive, multiplatform analysis of lung adenocarcinoma, with attention towards pathobiology and clinically actionable events.

Clinical samples and histopathologic data

We analysed tumour and matched normal material from 230 previously untreated lung adenocarcinoma patients who provided informed consent (Supplementary Table 1). All major histologic types of lung adenocarcinoma were represented: 5% lepidic, 33% acinar, 9% papillary, 14% micropapillary, 25% solid, 4% invasive mucinous, 0.4% colloid and 8% unclassifiable adenocarcinoma (Supplementary Fig. 1)[14]. Median follow-up was 19 months, and 163 patients were alive at the time of last follow-up. Eighty-one percent of patients reported past or present smoking. Supplementary Table 2 summarizes demographics. DNA, RNA and protein were extracted from specimens and quality-control assessments were performed as described previously[15]. Supplementary Table 3 summarizes molecular estimates of tumour cellularity[16].

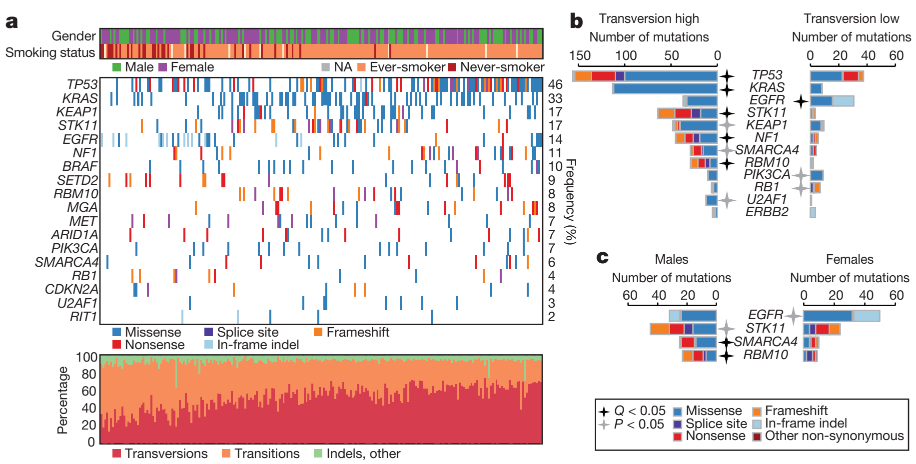


Figure 1 | Somatic mutations in lung adenocarcinoma. a, Co-mutation plot from whole exome sequencing of 230 lung adenocarcinomas. Data from TCGA samples were combined with previously published data[12] for statistical analysis. Co-mutation plot for all samples used in the statistical analysis ($n = 412$) can be found in Supplementary Fig. 2. Significant genes with a corrected P value less than 0.025 were identified using the MutSig2CV algorithm and are ranked in order of decreasing prevalence. **b**, **c**, The differential patterns of mutation between samples classified as transversion high and transversion low samples (**b**) or male and female patients (**c**) are shown for all samples used in the statistical analysis ($n = 412$). Stars indicate statistical significance using the Fisher's exact test (black stars: $q < 0.05$, grey stars: $P < 0.05$) and are adjacent to the sample set with the higher percentage of mutated samples.

*A list of authors and affiliations appears at the end of the paper.

Somatically acquired DNA alterations

We performed whole-exome sequencing (WES) on tumour and germline DNA, with a mean coverage of 97.6× and 95.8×, respectively, as performed previously[17]. The mean somatic mutation rate across the TCGA cohort was 8.87 mutations per megabase (Mb) of DNA (range: 0.5–48, median: 5.78). The non-synonymous mutation rate was 6.86 per Mb. MutSig2CV[18] identified significantly mutated genes among our 230 cases along with 182 similarly-sequenced, previously reported lung adenocarcinomas[12]. Analysis of these 412 tumour/normal pairs highlighted 18 statistically significant mutated genes (Fig. 1a shows co-mutation plot of TCGA samples ($n = 230$), Supplementary Fig. 2 shows co-mutation plot of all samples used in the statistical analysis ($n = 412$) and Supplementary Table 4 contains complete MutSig2CV results, which also appear on the TCGA Data Portal along with many associated data files (https://tcga-data.nci.nih.gov/docs/publications/luad_2014/). *TP53* was commonly mutated (46%). Mutations in *KRAS* (33%) were mutually exclusive with those in *EGFR* (14%). *BRAF* was also commonly mutated (10%), as were *PIK3CA* (7%), *MET* (7%) and the small GTPase gene, *RIT1* (2%). Mutations in tumour suppressor genes including *STK11* (17%), *KEAP1* (17%), *NF1* (11%), *RB1* (4%) and *CDKN2A* (4%) were observed. Mutations in chromatin modifying genes *SETD2* (9%), *ARID1A* (7%) and *SMARCA4* (6%) and the RNA splicing genes *RBM10* (8%) and *U2AF1* (3%) were also common. Recurrent mutations in the *MGA* gene (which encodes a Max-interacting protein on the MYC pathway[19]) occurred in 8% of samples. Loss-of-function (frameshift and nonsense) mutations in *MGA* were mutually exclusive with focal *MYC* amplification (Fisher's exact test $P = 0.04$), suggesting a hitherto unappreciated potential mechanism of MYC pathway activation. Coding single nucleotide variants and indel variants were verified by resequencing at a rate of 99% and 100%, respectively (Supplementary Fig. 3a, Supplementary Table 5). Tumour purity was not associated with the presence of false negatives identified in the validation data ($P = 0.31$; Supplementary Fig. 3b).

Past or present smoking associated with cytosine to adenine (C > A) nucleotide transversions as previously described both in individual genes and genome-wide[12,13]. C > A nucleotide transversion fraction showed two peaks; this fraction correlated with total mutation count ($R^2 = 0.30$) and inversely correlated with cytosine to thymine (C > T) transition frequency ($R^2 = 0.75$) (Supplementary Fig. 4). We classified each sample (Supplementary Methods) into one of two groups named transversion-high (TH, $n = 269$), and transversion-low (TL, $n = 144$). The transversion-high group was strongly associated with past or present smoking ($P < 2.2 \times 10^{-16}$), consistent with previous reports[13]. The transversion-high and transversion-low patient cohorts harboured different gene mutations. Whereas *KRAS* mutations were significantly enriched in the transversion-high cohort ($P = 2.1 \times 10^{-13}$), *EGFR* mutations were significantly enriched in the transversion-low group ($P = 3.3 \times 10^{-6}$). *PIK3CA* and *RB1* mutations were likewise enriched in transversion-low tumours ($P < 0.05$). Additionally, the transversion-low tumours were specifically enriched for in-frame insertions in *EGFR* and *ERBB2* (ref. 5) and for frameshift indels in *RB1* (Fig. 1b). *RB1* is commonly mutated in small-cell lung carcinoma (SCLC). We found *RB1* mutations in transversion-low adenocarcinomas were enriched for frameshift indels versus single nucleotide substitutions compared to SCLC ($P < 0.05$)[20,21] suggesting a mutational mechanism in transversion-low adenocarcinoma that is probably distinct from smoking in SCLC.

Gender is correlated with mutation patterns in lung adenocarcinoma[22]. Only a fraction of significantly mutated genes from the complete set reported in this study (Fig. 1a) were enriched in men or women (Fig. 1c). *EGFR* mutations were enriched in tumours from the female cohort ($P = 0.03$) whereas loss-of-function mutations within *RBM10*, an RNA-binding protein located on the X chromosome[23] were enriched in tumours from men ($P = 0.002$). When examining the transversion-high group, 16 out of 21 *RBM10* mutations were observed in males ($P = 0.003$, Fisher's exact test).

Somatic copy number alterations were very similar to those previously reported for lung adenocarcinoma[24] (Supplementary Fig. 5, Supplementary Table 6). Significant amplifications included *NKX2-1, TERT, MDM2, KRAS, EGFR, MET, CCNE1, CCND1, TERC* and *MECOM* (Supplementary Table 6), as previously described[24], 8q24 near *MYC*, and a novel peak containing *CCND3* (Supplementary Table 6). The *CDKN2A* locus was the most significant deletion (Supplementary Table 6). Supplementary Table 7 summarizes molecular and clinical characteristics by sample. Low-pass whole-genome sequencing on a subset ($n = 93$) of the samples revealed an average of 36 gene–gene and gene–inter-gene

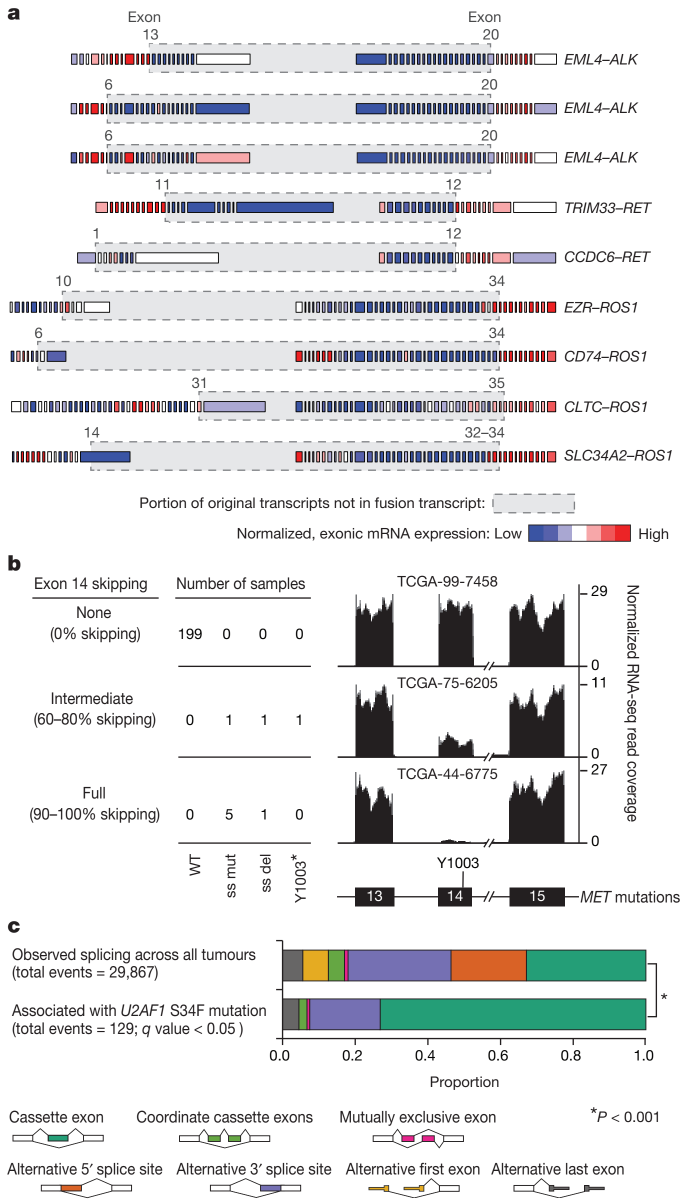


Figure 2 | Aberrant RNA transcripts in lung adenocarcinoma associated with somatic DNA translocation or mutation. a, Normalized exon level RNA expression across fusion gene partners. Grey boxes around genes mark the regions that are removed as a consequence of the fusion. Junction points of the fusion events are also listed in Supplementary Table 9. Exon numbers refer to reference transcripts listed in Supplementary Table 9. **b,** *MET* exon 14 skipping observed in the presence of exon 14 splice site mutation (ss mut), splice site deletion (ss del) or a Y1003* mutation. A total of 22 samples had insufficient coverage around exon 14 for quantification. The percentage skipping is (total expression minus exon 14 expression)/total expression. **c,** Significant differences in the frequency of 129 alternative splicing events in mRNA from tumours with *U2AF1* S34F tumours compared to *U2AF1* WT tumours (q value <0.05). Consistent with the function of U2AF1 in 3′ splice site recognition, most splicing differences involved cassette exon and alternative 3′ splice site events (chi-squared test, $P < 0.001$).

rearrangements per tumour. Chromothripsis[25] occurred in six of the 93 samples (6%) (Supplementary Fig. 6, Supplementary Table 8). Low-pass whole genome sequencing-detected rearrangements appear in Supplementary Table 9.

Description of aberrant RNA transcripts

Gene fusions, splice site mutations or mutations in genes encoding splicing factors promote or sustain the malignant phenotype by generating aberrant RNA transcripts. Combining DNA with mRNA sequencing enabled us to catalogue aberrant RNA transcripts and, in many cases, to identify the DNA-encoded mechanism for the aberration. Seventy-five per cent of somatic mutations identified by WES were present in the RNA transcriptome when the locus in question was expressed (minimum 5×) (Supplementary Fig. 7a) similar to prior analyses[15]. Previously identified fusions involving *ALK* (3/230 cases), *ROS1* (4/230) and *RET* (2/230) (Fig. 2a, Supplementary Table 10), all occurred in transversion-low tumours ($P = 1.85 \times 10^{-4}$, Fisher's exact test).

MET activation can occur by exon 14 skipping, which results in a stabilized protein[26]. Ten tumours had somatic *MET* DNA alterations with *MET* exon 14 skipping in RNA. In nine of these samples, a 5′ or 3′ splice site mutation or deletion was identified[27]. *MET* exon 14 skipping was also found in the setting of a *MET* Y1003* stop codon mutation (Fig. 2b, Supplementary Fig. 8a). The codon affected by the Y1003* mutation is predicted to disrupt multiple splicing enhancer sequences, but the mechanism of skipping remains unknown in this case.

S34F mutations in *U2AF1* have recently been reported in lung adenocarcinoma[12] but their contribution to oncogenesis remains unknown. Eight samples harboured *U2AF1*^{S34F}. We identified 129 splicing events strongly associated with *U2AF1*^{S34F} mutation, consistent with the role of U2AF1 in 3′-splice site selection[28]. Cassette exons and alternative 3′ splice sites were most commonly affected (Fig. 2c, Supplementary Table 11)[29]. Among these events, alternative splicing of the *CTNNB1* proto-oncogene was strongly associated with *U2AF1* mutations (Supplementary Fig. 8b). Thus, concurrent analysis of DNA and RNA enabled delineation of both *cis* and *trans* mechanisms governing RNA processing in lung adenocarcinoma.

Candidate driver genes

The receptor tyrosine kinase (RTK)/RAS/RAF pathway is frequently mutated in lung adenocarcinoma. Striking therapeutic responses are often achieved when mutant pathway components are successfully inhibited. Sixty-two per cent (143/230) of tumours harboured known activating mutations in known driver oncogenes, as defined by others[30]. Cancer-associated mutations in *KRAS* (32%, $n = 74$), *EGFR* (11%, $n = 26$) and *BRAF* (7%, $n = 16$) were common. Additional, previously uncharacterized *KRAS*, *EGFR* and *BRAF* mutations were observed, but were not classified as driver oncogenes for the purposes of our analyses (see Supplementary Fig. 9a for depiction of all mutations of known and unknown significance); explaining the differing mutation frequencies in each gene between this analysis and the overall mutational analysis described above. We also identified known activating *ERBB2* in-frame insertion and point mutations ($n = 5$)[6], as well as mutations in *MAP2K1* ($n = 2$), *NRAS* and *HRAS* ($n = 1$ each). RNA sequencing revealed the aforementioned *MET* exon 14 skipping ($n = 10$) and fusions involving *ROS1* ($n = 4$), *ALK* ($n = 3$) and *RET* ($n = 2$). We considered these tumours collectively as oncogene-positive, as they harboured a known activating RTK/RAS/RAF pathway somatic event. DNA amplification events were not considered to be driver events before the comparisons described below.

We sought to nominate previously unrecognized genomic events that might activate this critical pathway in the 38% of samples without a RTK/RAS/RAF oncogene mutation. Tumour cellularity did not differ between oncogene-negative and oncogene-positive samples (Supplementary Fig. 9b). Analysis of copy number alterations using GISTIC[31] identified unique focal *ERBB2* and *MET* amplifications in the oncogene-negative subset (Fig. 3a, Supplementary Table 6); amplifications in other wild-type proto-oncogenes, including *KRAS* and *EGFR*, were not significantly different between the two groups.

We next analysed WES data independently in the oncogene-negative and oncogene-positive subsets. We found that *TP53*, *KEAP1*, *NF1* and *RIT1* mutations were significantly enriched in oncogene-negative tumours ($P < 0.01$; Fig. 3b, Supplementary Table 12). *NF1* mutations have previously been reported in lung adenocarcinoma[11], but this is the first study, to our knowledge, capable of identifying all classes of loss-of-function

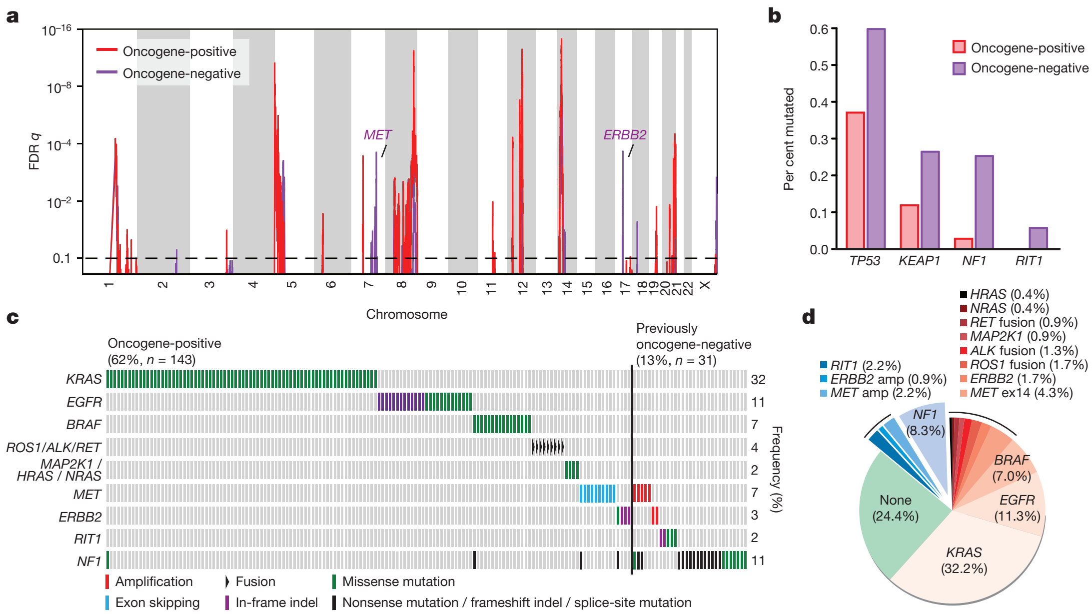


Figure 3 | Identification of novel candidate driver genes. **a,** GISTIC analysis of focal amplifications in oncogene-negative ($n = 87$) and oncogene-positive ($n = 143$) TCGA samples identifies focal gains of *MET* and *ERBB2* that are specific to the oncogene-negative set (purple). **b,** *TP53, KEAP1, NF1* and *RIT1* mutations are significantly enriched in samples otherwise lacking oncogene mutations (adjusted $P < 0.05$ by Fisher's exact test). **c,** Co-mutation plot of variants of known significance within the RTK/RAS/RAF pathway in lung adenocarcinoma. Not shown are the 63 tumours lacking an identifiable driver lesion. Only canonical driver events, as defined in Supplementary Fig. 9, and proposed driver events, are shown; hence not every alteration found is displayed. **d,** New candidate driver oncogenes (blue: 13% of cases) and known somatically activated drivers events (red: 63%) that activate the RTK/RAS/RAF pathway can be found in the majority of the 230 lung adenocarcinomas.

NF1 defects and to statistically demonstrate that *NF1* mutations, as well as *KEAP1* and *TP53* mutations are enriched in the oncogene-negative subset of lung adenocarcinomas (Fig. 3c). All *RIT1* mutations occurred in the oncogene-negative subset and clustered around residue Q79 (homologous to Q61 in the switch II region of *RAS* genes). These mutations transform NIH3T3 cells and activate MAPK and PI(3)K signalling[32], supporting a driver role for mutant *RIT1* in 2% of lung adenocarcinomas. This analysis increases the rate at which putative somatic lung adenocarcinoma driver events can be identified within the RTK/RAS/RAF pathway to 76% (Fig. 3d).

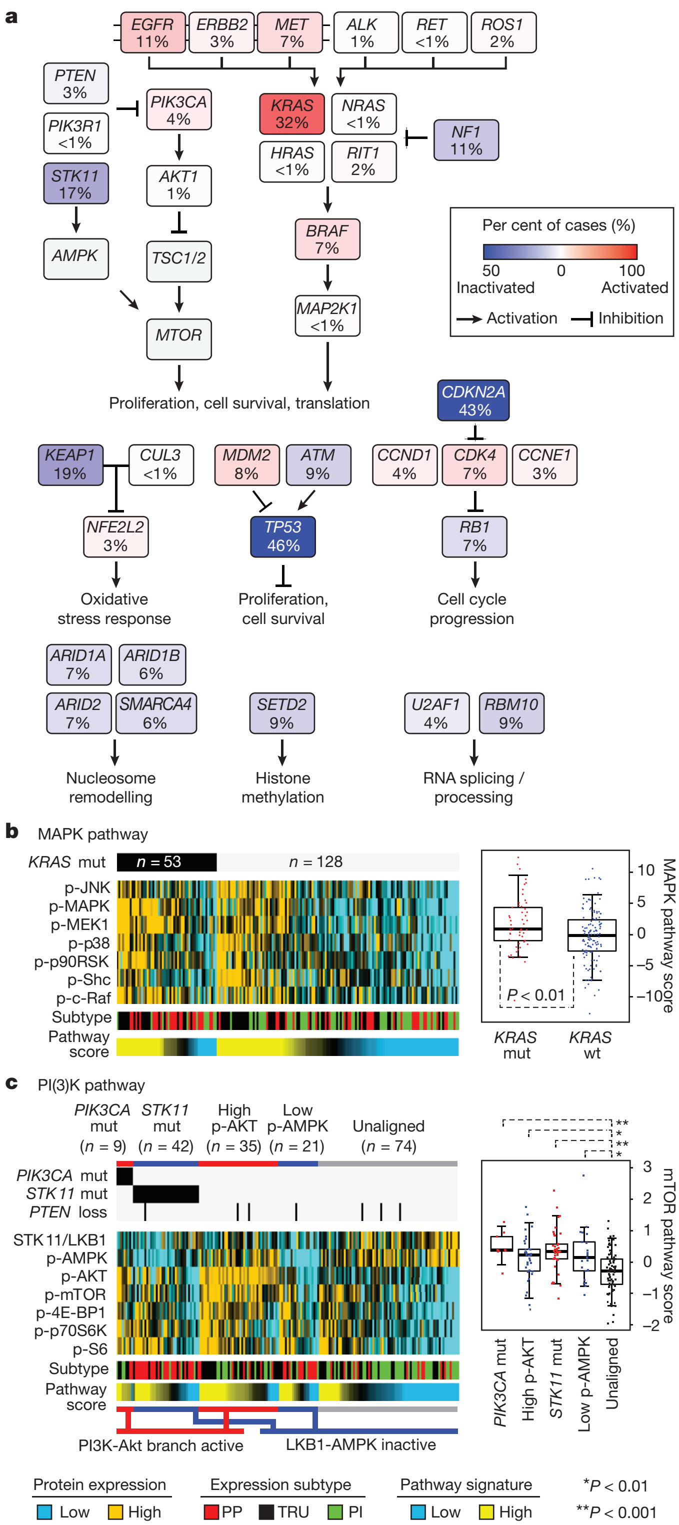


Figure 4 | Pathway alterations in lung adenocarcinoma. a, Somatic alterations involving key pathway components for RTK signalling, mTOR signalling, oxidative stress response, proliferation and cell cycle progression, nucleosome remodelling, histone methylation, and RNA splicing/processing. **b, c,** Proteomic analysis by RPPA ($n = 181$) P values by two-sided t-test. Box plots represent 5%, 25%, 75%, median, and 95%. PP, proximal proliferative; TRU, terminal respiratory unit; PI, proximal inflammatory. **c,** mTOR signalling may be activated, by either Akt (for example, via PI(3)K) or inactivation of AMPK (for example, via *STK11* loss). Tumours were separated into three main groups: those with PI(3)K-AKT activation, through either *PIK3CA* activating mutation or unknown mechanism (high p-AKT); those with LKB1-AMPK inactivation, through either *STK11* mutation or unknown mechanism with low levels of LKB1 and p-AMPK; and those showing none of the above features.

Recurrent alterations in key pathways

Recurrent aberrations in multiple key pathways and processes characterize lung adenocarcinoma (Fig. 4a). Among these were RTK/RAS/RAF pathway activation (76% of cases), PI(3)K-mTOR pathway activation (25%), p53 pathway alteration (63%), alteration of cell cycle regulators (64%, Supplementary Fig. 10), alteration of oxidative stress pathways (22%, Supplementary Fig. 11), and mutation of various chromatin and RNA splicing factors (49%).

We then examined the phenotypic sequelae of some key genomic events in the tumours in which they occurred. Reverse-phase protein arrays provided proteomic and phosphoproteomic phenotypic evidence of pathway activity. Antibodies on this platform are listed in Supplementary Table 13. This analysis suggested that DNA sequencing did not identify all samples with phosphoprotein evidence of activation of a given signalling pathway. For example, whereas *KRAS*-mutant lung adenocarcinomas had higher levels of phosphorylated MAPK than *KRAS* wild-type tumours had on average, many *KRAS* wild-type tumours displayed significant MAPK pathway activation (Fig. 4b, Supplementary Fig. 10). The multiple mechanisms by which lung adenocarcinomas achieve MAPK activation suggest additional, still undetected RTK/RAS/RAF pathway alterations. Similarly, we found significant activation of mTOR and its effectors (p70S6kinase, S6, 4E-BP1) in a substantial fraction of the tumours (Fig. 4c). Analysis of mutations in *PIK3CA* and *STK11*, STK11 protein levels, and AMPK and AKT phosphorylation[33] led to the identification of three major mTOR patterns in lung adenocarcinoma: (1) tumours with minimal or basal mTOR pathway activation, (2) tumours showing higher mTOR activity accompanied by either *STK11*-inactivating mutation or combined low STK11 expression and low AMPK activation and (3) tumours showing high mTOR activity accompanied by either phosphorylated AKT activation, *PIK3CA* mutation, or both. As with MAPK, many tumours lack an obvious underlying genomic alteration to explain their apparent mTOR activation.

Molecular subtypes of lung adenocarcinoma

Broad transcriptional and epigenetic profiling can reveal downstream consequences of driver mutations, provide clinically relevant classification and offer insight into tumours lacking clear drivers. Prior unsupervised analyses of lung adenocarcinoma gene expression have used varying nomenclature for transcriptional subtypes of the disease[34–37]. To coordinate naming of the transcriptional subtypes with the histopathological[38], anatomic and mutational classifications of lung adenocarcinoma, we propose an updated nomenclature: the terminal respiratory unit (TRU, formerly bronchioid), the proximal-inflammatory (PI, formerly squamoid), and the proximal-proliferative (PP, formerly magnoid)[39] transcriptional subtypes (Fig. 5a). Previously reported associations of expression signatures with pathways and clinical outcomes[34,36,39] were observed (Supplementary Fig. 7b) and integration with multi-analyte data revealed statistically significant genomic alterations associated with these transcriptional subtypes. The PP subtype was enriched for mutation of *KRAS*, along with inactivation of the *STK11* tumour suppressor gene by chromosomal loss, inactivating mutation, and reduced gene expression. In contrast, the PI subtype was characterized by solid histopathology and

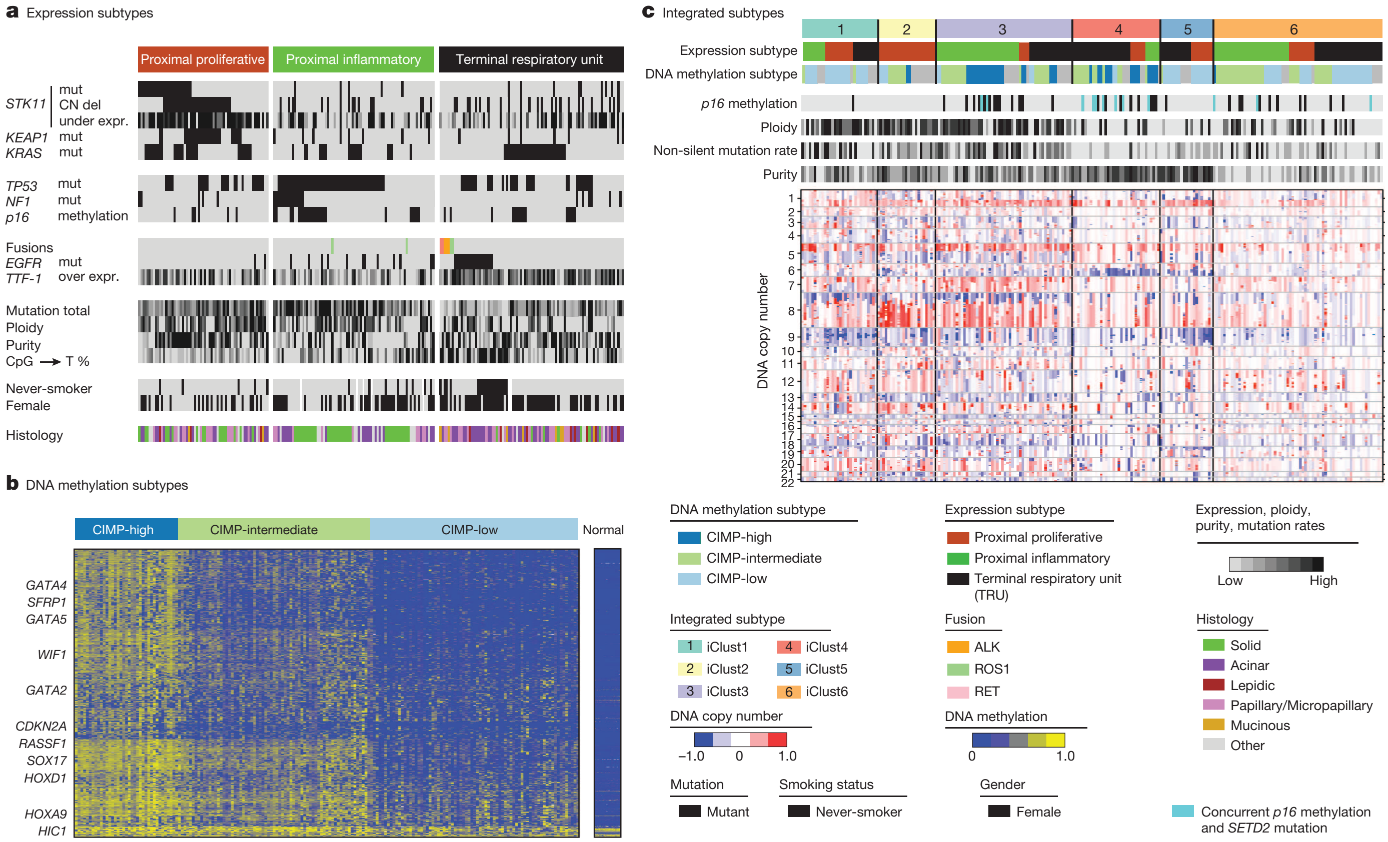


Figure 5 | Integrative analysis. a–c, Integrating unsupervised analyses of 230 lung adenocarcinomas reveals significant interactions between molecular subtypes. Tumours are displayed as columns, grouped by mRNA expression subtypes (**a**), DNA methylation subtypes (**b**), and integrated subtypes by iCluster analysis (**c**). All displayed features are significantly associated with subtypes depicted. The CIMP phenotype is defined by the most variable CpG island and promoter probes.

co-mutation of *NF1* and *TP53*. Finally, the TRU subtype harboured the majority of the *EGFR*-mutated tumours as well as the kinase fusion expressing tumours. TRU subtype membership was prognostically favourable, as seen previously[34] (Supplementary Fig. 7c). Finally, the subtypes exhibited different mutation rates, transition frequencies, genomic ploidy profiles, patterns of large-scale aberration, and differed in their association with smoking history (Fig. 5a). Unsupervised clustering of miRNA sequencing-derived or reverse phase protein array (RPPA)-derived data also revealed significant heterogeneity, partially overlapping with the mRNA-based subtypes, as demonstrated in Supplementary Figs 12 and 13.

Mutations in chromatin-modifying genes (for example, *SMARCA4*, *ARID1A* and *SETD2*) suggest a major role for chromatin maintenance in lung adenocarcinoma. To examine chromatin states in an unbiased manner, we selected the most variable DNA methylation-specific probes in CpG island promoter regions and clustered them by methylation intensity (Supplementary Table 14). This analysis divided samples into two distinct subsets: a significantly altered CpG island methylator phenotype-high (CIMP-H(igh)) cluster and a more normal-like CIMP-L(ow) group, with a third set of samples occupying an intermediate level of methylation at CIMP sites (Fig. 5b). Our results confirm a prior report[40] and provide additional insights into this epigenetic program. CIMP-H tumours often showed DNA hypermethylation of several key genes: *CDKN2A, GATA2, GATA4, GATA5, HIC1, HOXA9, HOXD13, RASSF1, SFRP1, SOX17* and *WIF1* among others (Supplementary Fig. 14). WNT pathway genes are significantly over-represented in this list (P value = 0.0015) suggesting that this is a key pathway with an important driving role within this subtype. *MYC* overexpression was significantly associated with the CIMP-H phenotype as well ($P = 0.003$).

Although we did not find significant correlations between global DNA methylation patterns and individual mutations in chromatin remodelling genes, there was an intriguing association between *SETD2* mutation and *CDKN2A* methylation. Tumours with low *CDKN2A* expression due to methylation (rather than due to mutation or deletion) had lower ploidy, fewer overall mutations (Fig. 5c) and were significantly enriched for *SETD2* mutation, suggesting an important role for this chromatin-modifying gene in the development of certain tumours.

Integrative clustering[41] of copy number, DNA methylation and mRNA expression data found six clusters (Fig. 5c). Tumour ploidy and mutation rate are higher in clusters 1–3 than in clusters 4–6. Clusters 1–3 frequently harbour *TP53* mutations and are enriched for the two proximal transcriptional subtypes. Fisher's combined probability tests revealed significant copy number associated gene expression changes on 3q in cluster one, 8q in cluster two, and chromosome 7 and 15q in cluster three (Supplementary Fig. 15). The low ploidy and low mutation rate clusters four and five contain many TRU samples, whereas tumours in cluster 6 have comparatively lower tumour cellularity, and few other distinguishing molecular features. Significant copy number-associated gene expression changes are observed on 6q in cluster four and 19p in cluster five. The CIMP-H tumours divided into a high ploidy, high mutation rate, proximal-inflammatory CIMP-H group (cluster 3) and a low ploidy, low mutation rate, TRU-associated CIMP-H group (cluster 4), suggesting that the CIMP phenotype in lung adenocarcinoma can occur in markedly different genomic and transcriptional contexts. Furthermore, cluster four is enriched for *CDKN2A* methylation and *SETD2* mutations, suggesting an interaction between somatic mutation of *SETD2* and deregulated chromatin maintenance in this subtype. Finally, cluster membership was significantly associated with mutations in *TP53*, *EGFR* and *STK11* (Supplementary Fig. 15, Supplementary Table 6).

Conclusions

We assessed the mutation profiles, structural rearrangements, copy number alterations, DNA methylation, mRNA, miRNA and protein expression

of 230 lung adenocarcinomas. In recent years, the treatment of lung adenocarcinoma has been advanced by the development of multiple therapies targeted against alterations in the RTK/RAS/RAF pathway. We nominate amplifications in *MET* and *ERBB2* as well as mutations of *NF1* and *RIT1* as driver events specifically in otherwise oncogene-negative lung adenocarcinomas. This analysis increases the fraction of lung adenocarcinoma cases with somatic evidence of RTK/RAS/RAF activation from 62% to 76%. While all lung adenocarcinomas may activate this pathway by some mechanism, only a subset show tonic pathway activation at the protein level, suggesting both diversity between tumours with seemingly similar activating events and as yet undescribed mechanisms of pathway activation. Therefore, the current study expands the range of possible targetable alterations within the RTK/RAS/RAF pathway in general and suggests increased implementation of MET and ERBB2/HER2 inhibitors in particular. Our discovery of inactivating mutations of *MGA* further underscores the importance of the MYC pathway in lung adenocarcinoma.

This study further implicates both chromatin modifications and splicing alterations in lung adenocarcinoma through the integration of DNA, transcriptome and methylome analysis. We identified alternative splicing due to both splicing factor mutations in *trans* and mutation of splice sites in *cis*, the latter leading to activation of the *MET* gene by exon 14 skipping. Cluster analysis separated tumours based on single-gene driver events as well as large-scale aberrations, emphasizing lung adenocarcinoma's molecular heterogeneity and combinatorial alterations, including the identification of coincident *SETD2* mutations and *CDKN2A* methylation in a subset of CIMP-H tumours, providing evidence of a somatic event associated with a genome-wide methylation phenotype. These studies provide new knowledge by illuminating modes of genomic alteration, highlighting previously unappreciated altered genes, and enabling further refinement in sub-classification for the improved personalization of treatment for this deadly disease.

METHODS SUMMARY

All specimens were obtained from patients with appropriate consent from the relevant institutional review board. DNA and RNA were collected from samples using the Allprep kit (Qiagen). We used standard approaches for capture and sequencing of exomes from tumour DNA and normal DNA[15] and whole-genome shotgun sequencing. Significantly mutated genes were identified by comparing them with expectation models based on the exact measured rates of specific sequence lesions[42]. GISTIC analysis of the circular-binary-segmented Affymetrix SNP 6.0 copy number data was used to identify recurrent amplification and deletion peaks[31]. Consensus clustering approaches were used to analyse mRNA, miRNA and methylation subtypes using previous approaches[15]. The publication web page is (https://tcga-data.nci.nih.gov/docs/publications/luad_2014/). Sequence files are in CGHub (https://cghub.ucsc.edu/).

Received 11 June 2013; accepted 22 April 2014.
Published online 9 July 2014.

1. Paez, J. G. *et al. EGFR* mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304,** 1497–1500 (2004).
2. Kwak, E. L. *et al.* Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N. Engl. J. Med.* **363,** 1693–1703 (2010).
3. Bergethon, K. *et al.* ROS1 rearrangements define a unique molecular class of lung cancers. *J. Clin Oncol.* **30,** 863–870 (2012).
4. Drilon, A. *et al.* Response to cabozantinib in patients with RET fusion-positive lung adenocarcinomas. *Cancer Discov.* **3,** 630–635 (2013).
5. Stephens, P. *et al.* Lung cancer: intragenic ERBB2 kinase mutations in tumours. *Nature* **431,** 525–526 (2004).
6. Takahashi, T. *et al.* p53: a frequent target for genetic abnormalities in lung cancer. *Science* **246,** 491–494 (1989).
7. Sanchez-Cespedes, M. *et al.* Inactivation of LKB1/STK11 is a common event in adenocarcinomas of the lung. *Cancer Res.* **62,** 3659–3662 (2002).
8. Shapiro, G. I. *et al.* Reciprocal Rb inactivation and p16^{INK4} expression in primary lung cancers and cell lines. *Cancer Res.* **55,** 505–509 (1995).
9. Singh, A. *et al.* Dysfunctional KEAP1–NRF2 interaction in non-small-cell lung cancer. *PLoS Med.* **3,** e420 (2006).
10. Medina, P. P. *et al.* Frequent *BRG1/SMARCA4*-inactivating mutations in human lung cancer cell lines. *Hum. Mutat.* **29,** 617–622 (2008).
11. Ding, L. *et al.* Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **455,** 1069–1075 (2008).
12. Imielinski, M. *et al.* Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150,** 1107–1120 (2012).
13. Govindan, R. *et al.* Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* **150,** 1121–1134 (2012).
14. Travis, W. D., Brambilla, E. & Riely, G. J. New pathologic classification of lung cancer: relevance for clinical practice and clinical trials. *J. Clin. Oncol.* **31,** 992–1001 (2013).
15. The Cancer Genome Atlas Research Network Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489,** 519–525 (2012).
16. Carter, S. L. *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nature Biotechnol.* **30,** 413–421 (2012).
17. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature Biotechnol.* **31,** 213–219 (2013).
18. Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505,** 495–501 (2014).
19. Hurlin, P. J., Steingrimsson, E., Copeland, N. G., Jenkins, N. A. & Eisenman, R. N. Mga, a dual-specificity transcription factor that interacts with Max and contains a T-domain DNA-binding motif. *EMBO J.* **18,** 7019–7028 (1999).
20. Peifer, M. *et al.* Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nature Genet.* **44,** 1104–1110 (2012).
21. Rudin, C. M. *et al.* Comprehensive genomic analysis identifies *SOX2* as a frequently amplified gene in small-cell lung cancer. *Nature Genet.* **44,** 1111–1116 (2012).
22. Tokumo, M. *et al.* The relationship between epidermal growth factor receptor mutations and clinicopathologic features in non-small cell lung cancers. *Clin. Cancer Res.* **11,** 1167–1173 (2005).
23. Coleman, M. P. *et al.* A novel gene, DXS8237E, lies within 20 kb upstream of UBE1 in Xp11.23 and has a different X inactivation status. *Genomics* **31,** 135–138 (1996).
24. Weir, B. A. *et al.* Characterizing the cancer genome in lung adenocarcinoma. *Nature* **450,** 893–898 (2007).
25. Stephens, P. J. *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144,** 27–40 (2011).
26. Kong-Beltran, M. *et al.* Somatic mutations lead to an oncogenic deletion of Met in lung cancer. *Cancer Res.* **66,** 283–289 (2006).
27. Seo, J. S. *et al.* The transcriptional landscape and mutational profile of lung adenocarcinoma. *Genome Res.* **22,** 2109–2119 (2012).
28. Wu, S., Romfo, C. M., Nilsen, T. W. & Green, M. R. Functional recognition of the 3′ splice site AG by the splicing factor U2AF^{35}. *Nature* **402,** 832–835 (1999).
29. Brooks, A. N. *et al.* A pan-cancer analysis of transcriptome changes associated with somatic mutations in *U2AF1* reveals commonly altered splicing events. *PLoS ONE* **9,** e87361 (2014).
30. Pao, W. & Hutchinson, K. E. Chipping away at the lung cancer genome. *Nature Med.* **18,** 349–351 (2012).
31. Beroukhim, R. *et al.* Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proc. Natl Acad. Sci. USA* **104,** 20007–20012 (2007).
32. Berger, A. H. *et al.* Oncogenic *RIT1* mutations in lung adenocarcinoma. *Oncogene* http://dx.doi.org/10.1038/onc.2013.581 (2014).
33. Creighton, C. J. *et al.* Proteomic and transcriptomic profiling reveals a link between the PI3K pathway and lower estrogen-receptor (ER) levels and activity in ER^{+} breast cancer. *Breast Cancer Res.* **12,** R40 (2010).
34. Wilkerson, M. D. *et al.* Differential pathogenesis of lung adenocarcinoma subtypes involving sequence mutations, copy number, chromosomal instability, and methylation. *PLoS ONE* **7,** e36530 (2012).
35. Beer, D. G. *et al.* Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nature Med.* **8,** 816–824 (2002).
36. Hayes, D. N. *et al.* Gene expression profiling reveals reproducible human lung adenocarcinoma subtypes in multiple independent patient cohorts. *J. Clin. Oncol.* **24,** 5079–5090 (2006).
37. Bhattacharjee, A. *et al.* Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc. Natl Acad. Sci. USA* **98,** 13790–13795 (2001).
38. Travis, W. D. *et al.* International association for the study of lung cancer/American Thoracic Society/European Respiratory Society international multidisciplinary classification of lung adenocarcinoma. *J. Thoracic Oncol.* **6,** 244–285 (2011).
39. Yatabe, Y., Mitsudomi, T. & Takahashi, T. TTF-1 expression in pulmonary adenocarcinomas. *Am. J. Surg. Pathol.* **26,** 767–773 (2002).
40. Shinjo, K. *et al.* Integrated analysis of genetic and epigenetic alterations reveals CpG island methylator phenotype associated with distinct clinical characters of lung adenocarcinoma. *Carcinogenesis* **33,** 1277–1285 (2012).
41. Mo, Q. *et al.* Pattern discovery and cancer gene identification in integrated cancer genomic data. *Proc. Natl Acad. Sci. USA* **110,** 4245–4250 (2013).
42. Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499,** 214–218 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements This study was supported by NIH grants: U24 CA126561, U24 CA126551, U24 CA126554, U24 CA126543, U24 CA126546, U24 CA137153, U24 CA126563, U24 CA126544, U24 CA143845, U24 CA143858, U24 CA144025, U24 CA143882, U24 CA143866, U24 CA143867, U24 CA143848, U24 CA143840, U24 CA143835, U24 CA143799, U24 CA143883, U24 CA143843, U54 HG003067, U54 HG003079 and U54 HG003273. We thank K. Guebert and L. Gaffney for assistance and C. Gunter for review.

Author Contributions The Cancer Genome Atlas Research Network contributed collectively to this study. Biospecimens were provided by the tissue source sites and processed by the biospecimen core resource. Data generation and analyses were performed by the genome sequencing centres, cancer genome characterization centres and genome data analysis centres. All data were released through the data coordinating centre. The National Cancer Institute and National Human Genome Research Institute project teams coordinated project activities. We also acknowledge the following TCGA investigators who made substantial contributions to the project: E. A. Collisson (manuscript coordinator); J. D. Campbell, J. Chmielecki, (analysis coordinators); C. Sougnez (data coordinator); J. D. Campbell, M. Rosenberg, W. Lee, J. Chmielecki, M. Ladanyi, and G. Getz (DNA sequence analysis); M. D. Wilkerson, A. N. Brooks, and D. N. Hayes (mRNA sequence analysis); L. Danilova and L. Cope (DNA methylation analysis); A. D. Cherniack (copy number analysis); M. D. Wilkerson and A. Hadjipanayis (translocations); N. Schultz, W. Lee, E. A. Collisson, A. H. Berger, J. Chmielecki, C. J. Creighton, L. A. Byers and M. Ladanyi (pathway analysis); A. Chu and A. G. Robertson (miRNA sequence analysis); W. Travis and D. A. Wigle (pathology and clinical expertise); L. A. Byers and G. B. Mills (reverse phase protein arrays); S. B. Baylin, R. Govindan and M. Meyerson (project chairs).

Author Information The primary and processed data used to generate the analyses presented here can be downloaded by registered users from The Cancer Genome Atlas at (https://tcga-data.nci.nih.gov/tcga/tcgaDownload.jsp). All of the primary sequence files are deposited in cgHub and all other data are deposited at the Data Coordinating Center (DCC) for public access (http://cancergenome.nih.gov/), (https://cghub.ucsc.edu/) and (https://tcga-data.nci.nih.gov/docs/publications/luad_2014/). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.M. (matthew_meyerson@dfci.harvard.edu).

The Cancer Genome Atlas Research Network

Disease analysis working group Eric A. Collisson[1], Joshua D. Campbell[2], Angela N. Brooks[2,3], Alice H. Berger[2], William Lee[4], Juliann Chmielecki[2], David G. Beer[5], Leslie Cope[6], Chad J. Creighton[7], Ludmila Danilova[6], Li Ding[8], Gad Getz[2,9,10], Peter S. Hammerman[2], D. Neil Hayes[11], Bryan Hernandez[2], James G. Herman[6], John V. Heymach[12], Igor Jurisica[13], Raju Kucherlapati[9], David Kwiatkowski[14], Marc Ladanyi[4], Gordon Robertson[15], Nikolaus Schultz[4], Ronglai Shen[4], Rileen Sinha[12], Carrie Sougnez[2], Ming-Sound Tsao[13], William D. Travis[4], John N. Weinstein[12], Dennis A. Wigle[16], Matthew D. Wilkerson[11], Andy Chu[15], Andrew D. Cherniack[2], Angela Hadjipanayis[9], Mara Rosenberg[2], Daniel J. Weisenberger[17], Peter W. Laird[17], Amie Radenbaugh[18], Singer Ma[18], Joshua M. Stuart[18], Lauren Averett Byers[12], Stephen B. Baylin[6], Ramaswamy Govindan[8], Matthew Meyerson[2,3]

Genome sequencing centres: The Eli & Edythe L. Broad Institute Mara Rosenberg[2], Stacey B. Gabriel[2], Kristian Cibulskis[2], Carrie Sougnez[2], Jaegil Kim[2], Chip Stewart[2], Lee Lichtenstein[2], Eric S. Lander[2,19] , Michael S. Lawrence[2], Getz[2,9,10]; **Washington University in St. Louis** Cyriac Kandoth[8], Robert Fulton[8], Lucinda L. Fulton[8], Michael D. McLellan[8], Richard K. Wilson[8], Kai Ye[8], Catrina C. Fronick[8], Christopher A. Maher[8], Christopher A. Miller[8], Michael C. Wendl[8], Christopher Cabanski[8], Li Ding[8], Elaine Mardis[8], Ramaswamy Govindan[8]; **Baylor College of Medicine** Chad J. Creighton[7], David Wheeler[7]

Genome characterization centres: Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Agency Miruna Balasundaram[15], Yaron S. N. Butterfield[15], Rebecca Carlsen[15], Andy Chu[15], Eric Chuah[15], Noreen Dhalla[15], Ranabir Guin[15], Carrie Hirst[15], Darlene Lee[15], Haiyan I. Li[15], Michael Mayo[15], Richard A. Moore[15], Andrew J. Mungall[15], Jacqueline E. Schein[15], Payal Sipahimalani[15], Angela Tam[15], Richard Varhol[15], A. Gordon Robertson[15], Natasja Wye[15], Nina Thiessen[15], Robert A. Holt[12], Steven J. M. Jones[15], Marco A. Marra[15]; **The Eli & Edythe L. Broad Institute** Joshua D. Campbell[2], Angela N. Brooks[2,3], Juliann Chmielecki[2], Marcin Imielinski[2,9,10], Robert C. Onofrio[2], Eran Hodis[9], Travis Zack[2], Carrie Sougnez[2], Elena Helman[2], Chandra Sekhar Pedamallu[2], Jill Mesirov[2], Andrew D. Cherniack[2], Gordon Saksena[2], Steven E. Schumacher[2], Scott L. Carter[2], Bryan Hernandez[2], Levi Garraway[2,3,9], Rameen Beroukhim[2,3,9], Stacey B. Gabriel[2], Gad Getz[2,9,10], Matthew Meyerson[2,3,9]; **Harvard Medical School/Brigham & Women's Hospital/MD Anderson Cancer Center** Angela Hadjipanayis[9,14], Semin Lee[9,14], Harshad S. Mahadeshwar[12], Angeliki Pantazi[9,14], Alexei Protopopov[12], Xiaojia Ren[9], Sahil Seth[12], Xingzhi Song[12], Jiabin Tang[12], Lixing Yang[9], Jianhua Zhang[12], Peng-Chieh Chen[9], Michael Parfenov[9,14], Andrew Wei Xu[9,14], Netty Santoso[9,14], Lynda Chin[12], Peter J. Park[9,14] & Raju Kucherlapati[9,14]; **University of North Carolina, Chapel Hill** Katherine A. Hoadley[11], J. Todd Auman[11], Shaowu Meng[11], Yan Shi[11], Elizabeth Buda[11], Scot Waring[11], Umadevi Veluvolu[11], Donghui Tan[11], Piotr A. Mieczkowski[11], Corbin D. Jones[11], Janae V. Simons[11], Matthew G. Soloway[11], Tom Bodenheimer[11], Stuart R. Jefferys[11], Jeffrey Roach[11], Alan P. Hoyle[11], Junyuan Wu[11], Saianand Balu[11], Darshan Singh[11], Jan F. Prins[11], J.S. Marron[11], Joel S. Parker[11], D. Neil Hayes[11], Charles M. Perou[11]; **University of Kentucky** Jinze Liu[20]; **The USC/JHU Epigenome Characterization Center** Leslie Cope[6], Ludmila Danilova[6], Daniel J. Weisenberger[17], Dennis T. Maglinte[17], Philip H. Lai[17], Moiz S. Bootwalla[17], David J. Van Den Berg[17], Timothy Triche Jr[17], Stephen B. Baylin[6], Peter W. Laird[17]

Genome data analysis centres: The Eli & Edythe L. Broad Institute Mara Rosenberg[2], Lynda Chin[12], Jianhua Zhang[12], Juok Cho[2], Daniel DiCara[2], David Heiman[2], Pei Lin[2], William Mallard[2], Douglas Voet[2], Hailei Zhang[2], Lihua Zou[2], Michael S. Noble[2], Michael S. Lawrence[2], Gordon Saksena[2], Nils Gehlenborg[2], Helga Thorvaldsdottir[2], Jill Mesirov[2], Marc-Danie Nazaire[2], Jim Robinson[2], Gad Getz[2,9,10]; **Memorial Sloan-Kettering Cancer Center** William Lee[4], B. Arman Aksoy[4], Giovanni Ciriello[4], Barry S. Taylor[1], Gideon Dresdner[4], Jianjiong Gao[4], Benjamin Gross[4], Venkatraman E. Seshan[4], Marc Ladanyi[4], Boris Reva[4], Rileen Sinha[4], S. Onur Sumer[4], Nils Weinhold[4], Nikolaus Schultz[4], Ronglai Shen[4], Chris Sander[4]; **University of California, Santa Cruz/Buck Institute** Sam Ng[18], Singer Ma[18], Jingchun Zhu[18], Amie Radenbaugh[18], Joshua M. Stuart[18], Christopher C. Benz[21], Christina Yau[21] & David Haussler[18,22]; **Oregon Health & Sciences University** Paul T. Spellman[23]; **University of North Carolina, Chapel Hill** Matthew D. Wilkerson[11], Joel S. Parker[11], Katherine A. Hoadley[11], Patrick K. Kimes[11], D. Neil Hayes[11], Charles M. Perou[11]; **The University of Texas MD Anderson Cancer Center** Bradley M. Broom[12], Jing Wang[12], Yiling Lu[12], Patrick Kwok Shing Ng[12], Lixia Diao[12], Lauren Averett Byers[12], Wenbin Liu[12], John V. Heymach[12], Christopher I. Amos[12], John N. Weinstein[12], Rehan Akbani[12], Gordon B. Mills[12]

Biospecimen core resource: International Genomics Consortium Erin Curley[24], Joseph Paulauskis[24], Kevin Lau[24], Scott Morris[24], Troy Shelton[24], David Mallery[24], Johanna Gardner[24], Robert Penny[24]

Tissue source sites: Analytical Biological Service, Inc. Charles Saller[25], Katherine Tarvin[25]; **Brigham & Women's Hospital** William G. Richards[14]; **University of Alabama at Birmingham** Robert Cerfolio[26], Ayesha Bryant[26]; **Cleveland Clinic**: Daniel P. Raymond[27], Nathan A. Pennell[27], Carol Farver[27]; **Christiana Care** Christine Czerwinski[28], Lori Huelsenbeck-Dill[28], Mary Iacocca[28], Nicholas Petrelli[28], Brenda Rabeno[28], Jennifer Brown[28], Thomas Bauer[28]; **Cureline** Oleg Dolzhanskiy[29], Olga Potapova[29], Daniil Rotin[29], Olga Voronina[29], Elena Nemirovich-Danchenko[29], Konstantin V. Fedosenko[29]; **Emory University** Anthony Gal[30], Madhusmita Behera[30], Suresh S. Ramalingam[30], Gabriel Sica[30]; **Fox Chase Cancer Center** Douglas Flieder[31], Jeff Boyd[31], JoEllen Weaver[31]; **ILSbio** Bernard Kohl[32], Dang Huy Quoc Thinh[32]; **Indiana University** George Sandusky[33]; **Indivumed** Hartmut Juhl[34]; **John Flynn Hospital** Edwina Duhig[35,36]; **Johns Hopkins University** Peter Illei[6], Edward Gabrielson[6], James Shin[6], Beverly Lee[6], Kristen Rogers[6], Dante Trusty[6], Malcolm V. Brock[6]; **Lahey Hospital & Medical Center** Christina Williamson[37], Eric Burks[37], Kimberly Rieger-Christ[37], Antonia Holway[37], Travis Sullivan[37]; **Mayo Clinic** Dennis A. Wigle[16], Michael K. Asiedu[16], Farhad Kosari[16]; **Memorial Sloan-Kettering Cancer Center** William D. Travis[4], Natasha Rekhtman[4], Maureen Zakowski[4], Valerie W. Rusch[4]; **NYU Langone Medical Center** Paul Zippile[38], James Suh[38], Harvey Pass[38], Chandra Goparaju[38], Yvonne Owusu-Sarpong[38]; **Ontario Tumour Bank** John M. S. Bartlett[39], Sugy Kodeeswaran[39], Jeremy Parfitt[39], Harmanjatinder Sekhon[39], Monique Albert[39]; **Penrose St. Francis Health Services** John Eckman[40], Jerome B. Myers[40]; **Roswell Park Cancer Institute** Richard Cheney[41], Carl Morrison[41], Carmelo Gaudioso[41]; **Rush University Medical Center** Jeffrey A. Borgia[42], Philip Bonomi[42], Mark Pool[42], Michael J. Liptay[42]; **St. Petersburg Academic University** Fedor Moiseenko[43], Irina Zaytseva[43]; **Thoraxklinik am Universitätsklinikum Heidelberg, Member of Biomaterial Bank Heidelberg (BMBH) & Biobank Platform of the German Centre for Lung Research (DZL)** Hendrik Dienemann[44], Michael Meister[44], Philipp A. Schnabel[45], Thomas R. Muley[44]; **University of Cologne** Martin Peifer[46]; **University of Miami** Carmen Gomez-Fernandez[47], Lynn Herbert[47], Sophie Egea[47]; **University of North Carolina** Mei Huang[11], Leigh B. Thorne[11], Lori Boice[11], Ashley Hill Salazar[11], William K. Funkhouser[11], W. Kimryn Rathmell[11]; **University of Pittsburgh** Rajiv Dhir[48], Samuel A. Yousem[48], Sanja Dacic[48], Frank Schneider[48], Jill M. Siegfried[48]; **The University of Texas MD Anderson Cancer Center** Richard Hajek[12]; **Washington University School of Medicine** Mark A. Watson[8], Sandra McDonald[8], Bryan Meyers[8]; **Queensland Thoracic Research Center** Belinda Clarke[35], Ian A. Yang[35], Kwun M. Fong[35], Lindy Hunter[35], Morgan Windsor[35], Rayleen V. Bowman[35]; **Center Hospitalier Universitaire Vaudois** Solange Peters[49], Igor Letovanec[49]; **Ziauddin University Hospital** Khurram Z. Khan[50]

Data Coordination Centre Mark A. Jensen[51], Eric E. Snyder[51], Deepak Srinivasan[51], Ari B. Kahn[51], Julien Baboud[51], David A. Pot[51]

Project team: National Cancer Institute Kenna R. Mills Shaw[52], Margi Sheth[52], Tanja Davidsen[52], John A. Demchok[52], Liming Yang[52], Zhining Wang[52], Roy Tarnuzzer[52], Jean Claude Zenklusen[52]; **National Human Genome Research Institute** Bradley A. Ozenberger[53], Heidi J. Sofia[53]

Expert pathology panel William D. Travis[4], Richard Cheney[41], Belinda Clarke[35], Sanja Dacic[48], Edwina Duhig[36,35], William K. Funkhouser[11], Peter Illei[6], Carol Farver[27], Natasha Rekhtman[4], Gabriel Sica[30], James Suh[38] & Ming-Sound Tsao[13]

[1]University of California San Francisco, San Francisco, California 94158, USA. [2]The Eli and Edythe L. Broad Institute, Cambridge, Massachusetts 02142, USA. [3]Dana Farber Cancer Institute, Boston, Massachusetts 02115, USA. [4]Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [5]University of Michigan, Ann Arbor, Michigan 48109, USA. [6]Johns Hopkins University, Baltimore, Maryland 21287, USA. [7]Baylor College of

Medicine, Houston, Texas 77030, USA. [8]Washington University, St. Louis, Missouri 63108, USA. [9]Harvard Medical School, Boston, Massachusetts 02115, USA. [10]Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [11]University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [12]University of Texas MD Anderson Cancer Center, Houston, Texas 77054, USA. [13]Princess Margaret Cancer Centre, Toronto, Ontario M5G 2M9, Canada. [14]Brigham and Women's Hospital Boston, Massachusetts 02115, USA. [15]BC Cancer Agency, Vancouver, British Columbia V5Z 4S6, Canada. [16]Mayo Clinic, Rochester, Minnesota 55905, USA. [17]University of Southern California, Los Angeles, California 90033, USA. [18]University of California Santa Cruz, Santa Cruz, California 95064, USA. [19]Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA. [20]University of Kentucky, Lexington, Kentucky 40515, USA. [21]Buck Institute for Age Research, Novato, California 94945, USA. [22]Howard Hughes Medical Institute, University of California Santa Cruz, Santa Cruz, California 95064, USA. [23]Oregon Health and Science University, Portland, Oregon 97239, USA. [24]International Genomics Consortium, Phoenix, Arizona 85004, USA. [25]Analytical Biological Services, Inc., Wilmington, Delaware 19801, USA. [26]University of Alabama at Birmingham, Birmingham, Alabama 35294, USA. [27]Cleveland Clinic, Cleveland, Ohio 44195, USA. [28]Christiana Care, Newark, Delaware 19713, USA. [29]Cureline, Inc., South San Francisco, California 94080, USA. [30]Emory University, Atlanta, Georgia 30322, USA. [31]Fox Chase Cancer Center, Philadelphia, Philadelphia 19111, USA. [32]ILSbio, Chestertown, Maryland 21620, USA. [33]Indiana University School of Medicine, Indianapolis, Indiana 46202, USA. [34]Individumed, Silver Spring, Maryland 20910, USA. [35]The Prince Charles Hospital and the University of Queensland Thoracic Research Center, Brisbane, 4032, Australia. [36]Sullivan Nicolaides Pathology & John Flynn Hospital, Tugun 4680, Australia. [37]Lahey Hospital and Medical Center, Burlington, Massachusetts 01805, USA. [38]NYU Langone Medical Center, New York, New York 10016, USA. [39]Ontario Tumour Bank, Ontario Institute for Cancer Research, Toronto, Ontario M5G 0A3, Canada. [40]Penrose St. Francis Health Services, Colorado Springs, Colorado 80907, USA. [41]Roswell Park Cancer Center, Buffalo, New York 14263, USA. [42]Rush University Medical Center, Chicago, Illinois 60612, USA. [43]St. Petersburg Academic University, St Petersburg 199034, Russia. [44]Thoraxklinik am Universitätsklinikum Heidelberg, 69126 Heidelberg, Germany. [45]University Heidelberg, 69120 Heidelberg, Germany. [46]University of Cologne, 50931 Cologne, Germany. [47]University of Miami, Sylvester Comprehensive Cancer Center, Miami, Florida 33136, USA. [48]University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA. [49]Center Hospitalier Universitaire Vaudois, Lausanne and European Thoracic Oncology Platform, CH-1011 Lausanne, Switzerland. [50]Ziauddin University Hospital, Karachi, 75300, Pakistan. [51]SRA International, Inc., Fairfax, Virginia 22033, USA. [52]National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. [53]National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA.

Lung cancer chemoprevention: current status and future prospects

Robert L. Keith and York E. Miller

Abstract | Lung cancer is the leading cause of cancer death worldwide, making it an attractive disease for chemoprevention. Although avoidance of tobacco use and smoking cessation will have the greatest impact on lung cancer development, chemoprevention could prove to be very effective, particularly in former smokers. Chemoprevention is the use of agents to reverse or inhibit carcinogenesis and has been successfully applied to other common malignancies. Despite prior studies in lung cancer chemoprevention failing to identify effective agents, we now have the ability to identify high-risk populations, and our understanding of lung tumour and premalignant biology continues to advance. There are distinct histological lesions that can be reproducibly graded as precursors of non-small-cell lung cancer and similar precursor lesions exist for adenocarcinoma. These premalignant lesions are being targeted by chemopreventive agents in current trials and will continue to be studied in the future. In addition, biomarkers that predict risk and response to targeted agents are being investigated and validated. In this Review, we discuss the principles of chemoprevention, data from preclinical models, completed clinical trials and observational studies, and describe new treatments for novel targeted pathways and future chemopreventive efforts.

Keith, R. L. & Miller, Y. E. *Nat. Rev. Clin. Oncol*. **10**, 334–343 (2013); published online 21 May 2013; doi:10.1038/nrclinonc.2013.64

Introduction

The term cancer chemoprevention was first introduced into the medical literature in 1976 by Michael Sporn and was defined to be the use of dietary or pharmaceutical interventions to slow or reverse the progression of premalignancy to invasive cancer. In this landmark publication,[1] Sporn discussed the evidence supporting the potential efficacy of retinoids in preventing lung cancer, a hypothesis that has subsequently undergone significant clinical testing with largely null or harmful outcomes in patients. For cancer chemoprevention to succeed, a high-risk population needs to be identifiable and agents that are both efficacious and associated with tolerable adverse effects must be available. For prevention of lung cancer, individuals at high risk can be readily identified using simple clinical features.[2,3] Unfortunately, no pharmacological or dietary intervention has been demonstrated to reduce lung cancer risk. Smoking cessation is currently the only known intervention effective in reducing the risk of lung cancer.[4]

Lung cancer is the leading cause of cancer death in the world, with an estimated 1,387,400 deaths in 2011,[5] making it an attractive target disease for cancer chemoprevention. In the USA, over 200,000 lung cancer diagnoses were made in 2012 and the 5-year overall survival for patients is a dismal 16%.[6] Although gradual improvements in survival have been achieved, this advance has not matched those seen for other common malignancies, such as breast, prostate and colon cancer, partly as a result of lung cancer patients often presenting at an advanced stage where surgical cure is no longer feasible. Early detection is critical for improving outcomes, and lung cancer screening using low-dose CT scans has been shown to be effective in reducing mortality by 20%.[7] Although this is a major advance in preventing deaths from lung cancer, even with the widespread introduction of CT screening, overall survival remains low, with less than 20% of patients living beyond 5 years. Populations at high risk for lung cancer, with an annual incidence of up to a 2%, can be readily identified using easily obtainable information, including smoking history, previous history of a tobacco-induced cancer, pulmonary function, and family history.[3,8,9] More sophisticated risk assessment models have been developed, which include capacity for DNA repair, but these models do not add greatly to the previously described models.[9] For breast cancer, a predicted annual incidence of 0.3% is an indication for consideration of chemoprevention treatment, so the potential population for lung cancer chemoprevention is comparatively large.

The vast majority of lung cancers diagnosed in the USA (85–90%) are associated with exposure to tobacco smoke.[10] Therefore, prevention of smoking initiation is clearly the most effective intervention that can be applied to reduce the burden of lung cancer. Smoking cessation has been demonstrated to decrease lung cancer and prolong life.[4,11] Most recently, this was shown in the Million Women Study in the UK, where female smokers who continued to smoke beyond the age of 40 lost at

Division of Pulmonary Sciences and Critical Care Medicine, Department of Medicine, VA Eastern Colorado Healthcare System, University of Colorado Denver School of Medicine, 1055 Clermont Street, Box 151, Denver, CO 80220, USA (**R. L. Keith, Y. E. Miller**).

Correspondence to: R. L. Keith robert.keith@ucdenver.edu

Competing interests
R. L. Keith declares an association with the following companies: Boehringer-Ingelheim, Pfizer. R. L. Keith and Y. E. Miller declare an association with Bayer–Schering. See the article online for full details of the relationship.

least 10 years of lifespan.[12] In contrast to other smoking-related diseases, former heavy smokers retain a significant risk of developing lung cancer years after smoking cessation. Currently in the USA, more than 50% of lung cancer occurs in former smokers.[13,14] Studies have shown that response to chemopreventive agents may differ between current and former smokers, with a more favourable response in the latter.[15,16] Therefore, limiting chemoprevention clinical trial enrolment to former smokers or analysis of the two groups separately remains a reasonable strategy.

In this Review, we provide an overview of previous clinical trials that assessed lung cancer prevention as the primary goal. We will also discuss strategies for identifying new agents for lung cancer prevention, including modulation of intermediate end point biomarkers, preclinical models and observational studies. Finally, we will summarize our vision of novel ways to move the field forward.

Key points

- The annual risk for lung cancer in patient populations readily identifiable using simple clinical and demographic characteristics approaches 2%; therefore, the potential for chemoprevention in this common cancer is high
- No chemopreventive agents have been shown to be efficacious for lung cancer, despite numerous leads from observational studies
- Smoking cessation remains the only intervention proven to reduce lung cancer risk
- Lung cancer is a heterogeneous malignancy from a mutational standpoint; focusing on frequently altered pathways may be a promising strategy because molecular targeting of chemoprevention to specific mutations is challenging
- Chemoprevention targeted to phenotypes expressing specific carcinogenic influences, including inflammation, angiogenesis, hypoxia and epithelial differentiation, appears most likely to succeedphenotypes

Human pulmonary carcinogenesis

Lung cancer is currently divided into histological classifications of small-cell carcinoma and the non-small-cell carcinoma types of adenocarcinoma, squamous-cell carcinoma, and large-cell undifferentiated carcinoma. Adenocarcinoma and squamous-cell lung cancer make up the vast majority of lung cancer cases, and are the only cell types for which premalignant histology has been well characterized. For both of these subtypes of lung cancer, multistep carcinogenesis involving genetic and epigenetic alterations in pulmonary epithelial cells has been demonstrated.[17] However, a consistent progression of events has not been clearly established.

For squamous-cell lung cancer, progression through a series of histological changes, which can be sampled by bronchoscopy, has been described. These changes, classified by the WHO, include reserve cell hyperplasia, squamous metaplasia, mild, moderate and severe dysplasia, and carcinoma *in situ* (Figure 1).[18] The lesions can be reproducibly graded,[18] but the risk of progression of each lesion to invasive squamous-cell lung cancer has not been well established because this process is difficult to study and, therefore, remains controversial. Several studies have suggested that assessment of the accumulation of genetic alterations—assessed by fluorescence *in situ* hybridization (FISH),[19] PCR-based loss of heterozygosity,[20] comparative genomic hybridization[21] or sequence analysis[22]—might provide important prognostic information beyond that of the level of histological dysplasia.[23] Amplification of genes encoded on chromosome 3q26,[21] including *SOX2*, *PIK3CA* and *TP63*, has been identified as an early event in the development of squamous-cell lung cancer, with *SOX2* amplification occurring before the other genetic changes.[24]

The premalignant biology of adenocarcinoma is more difficult to investigate than that of squamous-cell lung cancer, since lesions are only accessible by surgical resection or percutaneous needle biopsy. Adenocarcinoma seems to be preceded by a premalignant lesion (atypical adenomatous hyperplasia) and the preinvasive lesion (adenocarcinoma *in situ,* formerly known as bronchoalveolar carcinoma), which progresses to invasive adenocarcinoma.[25] In some cases, specific *EGFR* or *KRAS* mutations might precede the development of invasive squamous cell or adenocarcinoma.[26–28]

In 1953, Slaughter *et al.*[29] introduced the term 'field cancerization' to describe the frequent occurrence of widespread premalignant histological changes and second primary neoplasms in smokers with oral cancer. It is now widely accepted that field cancerization also occurs in the lower respiratory tract during pulmonary carcinogenesis, with the molecular correlates of widely dispersed epithelial cells harbouring mutations in either *TP53* or *EGFR* associated with squamous-cell or adenocarcinoma premalignancy, respectively.[22,26] EGFR or p53 protein overexpression has been reported in premalignant dysplasia,[27,30] and expression is most pronounced in more-advanced lesions (that is, severe dysplasia and carcinoma *in situ*).

Chemoprevention

Carcinogenesis is a complex process, involving carcinogen exposure and activation, DNA adduct formation, inflammation, oxidative stress, mutation and epigenetic alterations, which lead to the acquisition of the hallmarks of cancer.[31] These hallmarks include: sustained proliferative signalling; growth suppression evasion; cell death resistance; replicative immortality; angiogenesis; invasion or metastasis; reprogrammed energy metabolism, and immune evasion. All these hallmarks arise as a result of genomic instability. Chemoprevention efforts have been directed at all of the carcinogenic processes, as well as at many of these intrinsic features of cancer (Figure 1). Both adenocarcinoma and squamous-cell carcinoma of the lung are genetically complex and heterogeneous, making chemoprevention with an agent targeted to specific driver mutations unlikely to be widely effective, except perhaps in the setting of definable premalignant lesions (ground glass opacities or endobronchial dysplasias) with established mutations in targetable proteins.[32,33] Loss of function of the p53 tumour-suppressor protein might be the most common mutation in squamous cell and adenocarcinoma, but approaches to restore p53 function have yet to be therapeutically translated. We speculate that the most broadly effective chemoprevention approaches will

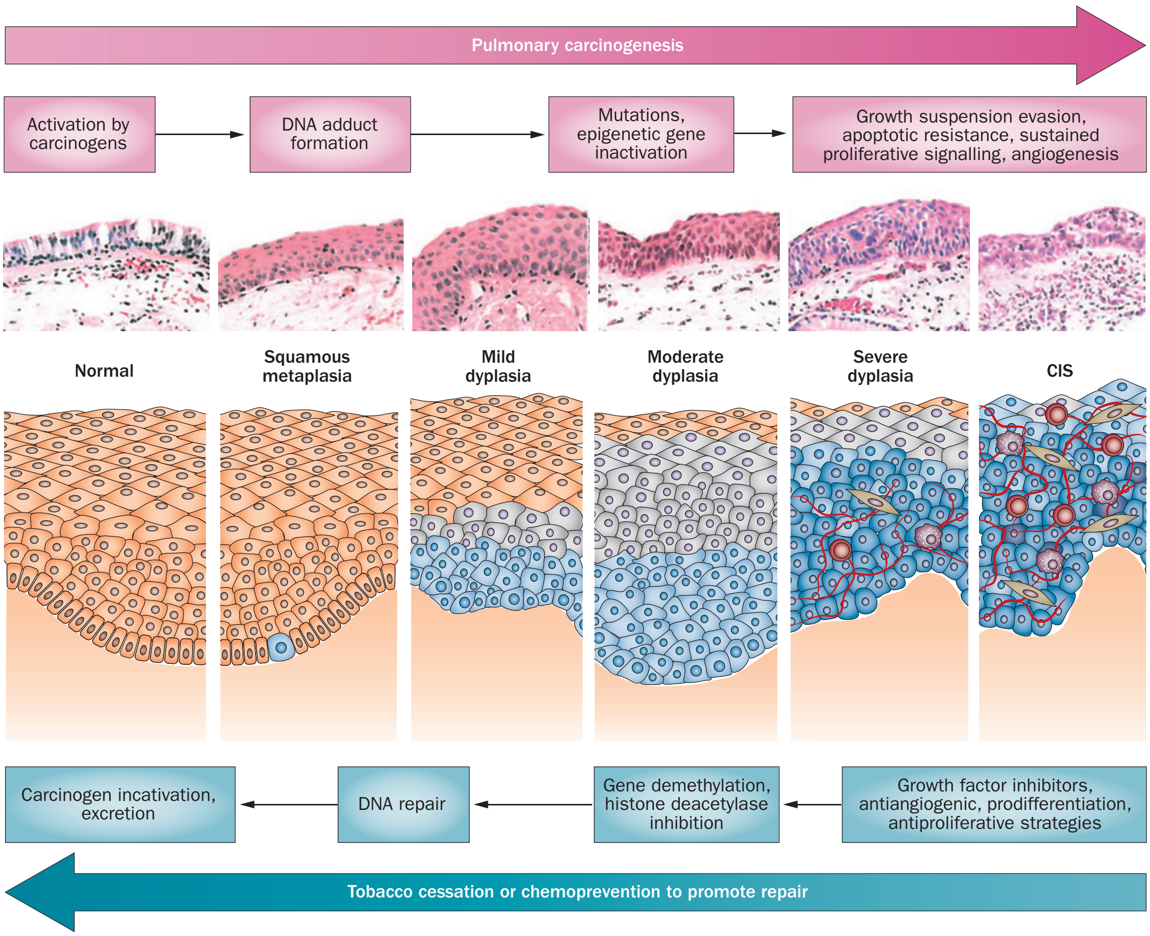


Figure 1 | Potential tobacco smoke induced carcinogenic processes for chemopreventive intervention. Hallmarks in the development of squamous-cell lung cancer as the bronchial epithelium proceeds through pathological stages in the progression to CIS. Tobacco cessation and chemopreventive agents can promote repair and block progression. Abbreviation: CIS, carcinoma *in situ*.

target general processes, such as suppression of inflammation, interference with autocrine or paracrine growth stimulation, restoration of epithelial differentiation and polarity, augmentation of apoptosis, improved immune surveillance, and suppression of invasion or angiogenesis.

Phase III chemoprevention trials

Chemoprevention efforts fall into three distinct subgroups: primary, secondary and tertiary. Primary chemoprevention involves patients at increased risk, but without a previous history of cancer. Secondary chemoprevention studies enroll individuals with increased risk and evidence of premalignancy. For lung cancer, this usually refers to sputum cytological atypia and/or endobronchial dysplasia. Recent studies have also focused on individuals with ground glass opacities on CT scan suggestive of adenocarcinoma *in situ* or atypical adenomatous hyperplasia.[34–36] Tertiary chemoprevention trials have the end point of a second primary tumour in individuals with a previous tobacco-induced aerodigestive cancer. A number of phase III chemoprevention trials, including primary, secondary and tertiary studies, have been reported in the past two decades (Table 1). Unfortunately, the results of phase III lung cancer chemoprevention trials can be summarized succinctly: aspirin,[37–39] retinyl palmitate,[40] 13-*cis*-retinoic acid,[41] vitamin E,[42] a multivitamin and mineral supplement,[43] and selenium[44] are all ineffective; beta-carotene seems to be harmful in current smokers.[16]

As the discouraging results of phase III lung cancer chemoprevention trials have accumulated, investigators and funding agencies have adopted increasingly stringent criteria for assessing a treatment in a phase III trial.[45] Although all criteria cannot be applied to every putative agent, evidence supporting efficacy of a chemopreventive intervention ideally should be derived from multiple sources, including mechanistic, preclinical, observational and phase II studies with intermediate or surrogate end points. In addition, ideal chemopreventive agents would be well tolerated, inexpensive, and perhaps also treat comorbid disease (such as chronic obstructive pulmonary disease [COPD], diabetes mellitus, pulmonary hypertension, or atherosclerosis) in high-risk individuals. Currently, no interventions other than smoking cessation have been shown to reduce lung cancer risk, so we believe that efficacy should take precedence over these ancillary characteristics as attempts are made to discover ways to reduce lung cancer risk.

Early phase trials

Although phase III clinical trials represent the definitive means of demonstrating efficacy in terms of reducing cancer incidence and mortality, phase II cancer prevention trials rely on intermediate end points that are meant to predict these outcomes. The terms 'surrogate end point' and 'intermediate end point' are not synonymous. A surrogate end point is obtained earlier, potentially less invasively, and at lower cost than a true end point, but serves as a substitute for the true outcome. An intermediate end point should be integrally involved in the disease process (carcinogenesis in this case), and expression should differ between normal and at-risk subjects and correlate with disease course.[46]

Owing to the lack of an effective chemoprevention agent for lung cancer (Table 1), no intermediate end point biomarkers have been validated as surrogate end points according to the Prentice criteria.[47] Therefore, intermediate end points for phase II chemoprevention trials are largely chosen on the basis of biological plausibility. Premalignant histology is the most-widely accepted intermediate end point, but has several weaknesses. The WHO classification focuses largely on squamous-cell precursor lesions, which are accessible to bronchoscopic biopsy.[18] Whether an effect on these squamous precursor lesions might carry over to premalignant adenocarcinoma lesions is speculative. The complexity of branching airway anatomy and small calibre of higher division airways makes bronchoscopic inspection of the entire epithelial surface impossible. It is extremely rare to biopsy a premalignant dysplasia that subsequently develops into an invasive squamous-cell carcinoma, so accurately predicting the progression of premalignant lesions (except possibly for carcinoma *in situ)* is infrequent and controversial. Although endobronchial histology has a poor correlation with lung cancer risk, the addition of biomarkers of genetic alteration (including chromosomal aneusomy detected by FISH, PCR-detected loss of heterozygosity or gene-copy number changes detected by comparative genomic hybridization or single nucleotide polymorphism array) are promising.[19–22] Proliferation index, most commonly measured by Ki-67 immunostaining, is a similarly plausible, but unvalidated intermediate end point biomarker.[48] A number of novel biomarkers have been developed that might be validated and could be used in phase II chemoprevention trials. These include a transcriptomic signature derived from endobronchial or nasal brushings,[49,50] biopsy proteomics,[51] serum proteomics,[52] and the analysis of volatile organic compounds in exhaled breath.[53]

Several phase II chemoprevention trials have been completed (Table 2), although few have met their primary end point. Two negative phase II trials of 13-*cis*-retinoic acid, with histological end points,[54,55] have been published and the results are consistent with the negative phase III trial of 13-*cis*-retinoic acid in individuals with a previous lung cancer diagnosis.[41] This consistency at least suggests that phase II trials might help identify potential chemopreventive agents that should not progress to phase III testing. To date, the only trials to have met the primary end point are those assessing iloprost (a drug to treat pulmonary arterial hypertension that produced histological improvement[56]) and celecoxib (decreased Ki-67 levels in two trials[57,58]). In the iloprost trial, Ki-67 level was a secondary end point and was not decreased by iloprost treatment. However, there were considerable differences between the celecoxib and iloprost trials. In the celecoxib trials, dysplasia was either extremely infrequent or not reported, and the iloprost trial evaluated a cohort with an approximately 70% incidence of dysplasia. Expression of Ki-67 was assessed largely in the normal bronchial epithelium in the celecoxib trials and in more-advanced lesions in the iloprost trial. As no intermediate end points have yet been validated to be predictive of chemopreventive efficacy, it is not clear if either the modulation of histology or Ki-67 index is informative.

Table 1 | Phase III chemoprevention trials

Intervention	Chemoprevention setting	*n*	Results
Aspirin	Primary	5,139 22,071 39,876	Negative[39] Negative[37] Negative[38]
Beta carotene	Primary	29,133 22,071	Harmful (RR = 1.18)[42] Negative[15]
Beta carotene and retinol	Primary	18,314	Harmful (RR = 1.28)[122]
Multivitamins and minerals	Primary	29,584	Negative[43]
Vitamin E	Primary	29,133	Negative[42]
Retinyl palmitate	Tertiary	2,592	Negative[40]
13-*cis*-retinoic acid	Tertiary	1,166	Negative[41]
N-acetyl cysteine	Tertiary	2,592	Negative[40]
Selenium	Tertiary	1,772	Negative[44]

Abbreviation: RR, relative risk.

Although most phase II trials have assessed histological or proliferative intermediate end points in the central airways, three trials evaluated CT-detected pulmonary nodules in response to inhaled corticosteroids.[34–36] CT-detected nodules were not the primary end point in two of these trials, but were included in exploratory analyses that indicated efficacy of inhaled corticosteroids, and led to a third trial with nodules as the primary end point. The latter trial,[34] was negative in that nodules were not decreased by inhaled corticosteroids. This potentially exciting new end point might represent atypical adenomatous hyperplasia or adenocarcinoma *in situ,* which presents as ground glass opacities on thoracic CT scan. However, this end point lacks specificity as other pulmonary conditions, including infection, organizing pneumonia, hypersensitivity pneumonitis, and desquamative interstitial pneumonia, can also cause ground glass opacities.[59] Short-term trials assessing the response of the ground glass component of semi-solid pulmonary nodules that will undergo resection might address this lack of specificity, but we are not aware of any such trials that are ongoing or planned.

The phase I trial that assessed myoinositol had no control group, but gave promising results when compared

Table 2 | Early phase intermediate end point chemoprevention trials

Intervention	End point	n	Outcome
13-*cis*-retinoic acid	Metaplasia Dysplasia	40 100	Negative[55] Negative[54]
Fenretinide	Metaplasia	82	Negative[123]
Etretinate	Sputum atypia	150	Negative[124]
Beta carotene	Sputum atypia	1,067	Negative[125]
Vitamin B12/folate	Sputum atypia	73	Negative[126]
Budesonide	Dysplasia Nodule size	112 202	Negative for primary end point; nodules decreased in treatment group[35] Negative[34]
Fluticasone	Nodule size and number	201	Negative[36]
Anethole dithiolethione	New dysplastic lesions	101	Negative for primary end point; rate of worsening lower in treatment group[127]
Iloprost	Dysplasia	152	Positive in former smokers only (improvement in airway endobronchial histology)[55]
Celecoxib	Ki-67 Ki-67	204 101	Positive (decreased Ki-67 labelling in former smokers)[56] Positive (decreased Ki-67 labelling in former smokers)[57]
Myoinositol	Dysplasia	26	Promising, was a phase I trial with historical controls[59]

to historical controls.[60] Of interest, gene-expression profiling of bronchial brushings identified increased expression of PI3K in association with lung cancer or dysplastic lesions. [61] Myoinositol inhibits PI3K, and elevated expression of genes involved in PI3K signalling identified patients who responded to myoinositol. These data indicate that the personalization of chemoprevention may be possible, as is the case for treating lung cancer with targeted agents.[61] Results of an ongoing randomized controlled phase II trial of myoinositol will be of great interest.

Preclinical studies

Similar to other common cancers, animal models of non-small-cell lung cancer have been extensively used to study lung carcinogenesis and perform preclinical studies involving chemopreventive agents. Multiple, well-characterized models of murine adenocarcinoma are available, including initiator-promoter carcinogenesis,[62] mutant *KRAS*[63] or *EGFR*[64] and the use of complete carcinogens.[65–67] Tobacco smoke is a mouse lung carcinogen and can reproducibly induce pulmonary adenocarcinomas, but it is a labour-intensive and expensive model.[68] More recently, a chemically induced squamous-cell lung carcinoma model has been described, and histopathological analysis of serial lung sections in this model revealed a range of lung pathology, including squamous-cell carcinoma, carcinoma *in situ,* and varying levels of bronchial dysplasia.[69] Immunohistochemical studies on the premalignant lesions show staining that corresponds to analogous human lesions.[70] Gene expression similarities between human and murine adenocarcinoma have been described;[71] these comparisons have not yet been made for the squamous-cell carcinoma model.

Eicosanoids represent a large family of bioactive lipid molecules that signal through autocrine and paracrine pathways and have been implicated in cancer initiation, progression and metastasis (Figure 2).[72] Prostaglandin I_2 (PGI_2, prostacyclin) is a member of the eicosanoid family that has anti-inflammatory, antiproliferative, anti-metastatic, and potent chemopreventive properties.[73] The chemopreventive effects of PGI_2, or its analogue iloprost, are independent of the canonical cell surface receptor (prostacyclin receptor), but rather involve the activation of the nuclear receptor peroxisome proliferator-activated receptor gamma (PPARγ).[74] PPARγ ligands inhibit the growth of lung cancer cell lines *in vitro* and in xenograft models, resulting in decreased proliferation, increased apoptosis, and promotion of differentiation (which may allow for the use of targeted therapeutics).[75,76] Furthermore, multiple animal carcinogenesis studies have demonstrated that PPARγ ligand treatment inhibits lung tumour development and can induce apoptosis.[77,78] The observed tumour inhibition occurs in the setting of mutated *TP53*, thereby mirroring the most-common human genetic abnormality found in lung cancer. The addition of the PPARγ ligand pioglitazone to the inhaled steroid budesonide further improved the efficacy associated with either agent alone and was shown to decrease tumour load by 90% in animals exposed to the cigarette smoke carcinogen benzo(a)pyrene.[78,79] Mice overexpressing lung-specific PPARγ developed 70% fewer tumours than wild-type controls when exposed to carcinogens.[74]

Chemopreventive interventions have been assessed in murine preclinical models. Inhaled and systemic glucocorticoids,[80] myoinositol,[80] overexpression of PGI_2,[81] dietary administration of iloprost,[74] overexpression of PPARγ,[74] dietary administration of pioglitazone[78] and the VEGF inhibitor vandetanib,[82] as well as the anti-oestrogen fulvestrant, have all proved efficacious in murine models.[83] The effect of COX inhibitors on lung cancer prevention has also been tested in murine models. The COX inhibitor indomethacin was found to be effective,[84] but the COX-2 selective inhibitor celecoxib was not.[85] Additionally, rexinoids, agents that activate nuclear retinoid X receptors, and triterpenoids[86] have been shown to prevent murine adenocarcinoma, both alone and when used in combination.[87,88]

Observational studies

Observational studies on the relationship of aspirin use and lung cancer are inconsistent; any positive reports are at odds with the results of randomized controlled studies including the Physicians' Health Study[37] and Women's Health Study.[38] Recently, a meta-analysis of case-control and cohort studies of aspirin in the prevention of vascular events was published in which the association between aspirin use and the incidence of several cancers was analysed. Regular use of aspirin was associated with a reduced risk of colorectal cancer (odds ratio = 0.62, 95% CI 0.58–0.67; P < 0.0001), but only a trend towards protection from lung cancer was identified.[89] Another meta-analysis of 15 observational studies (six case-control

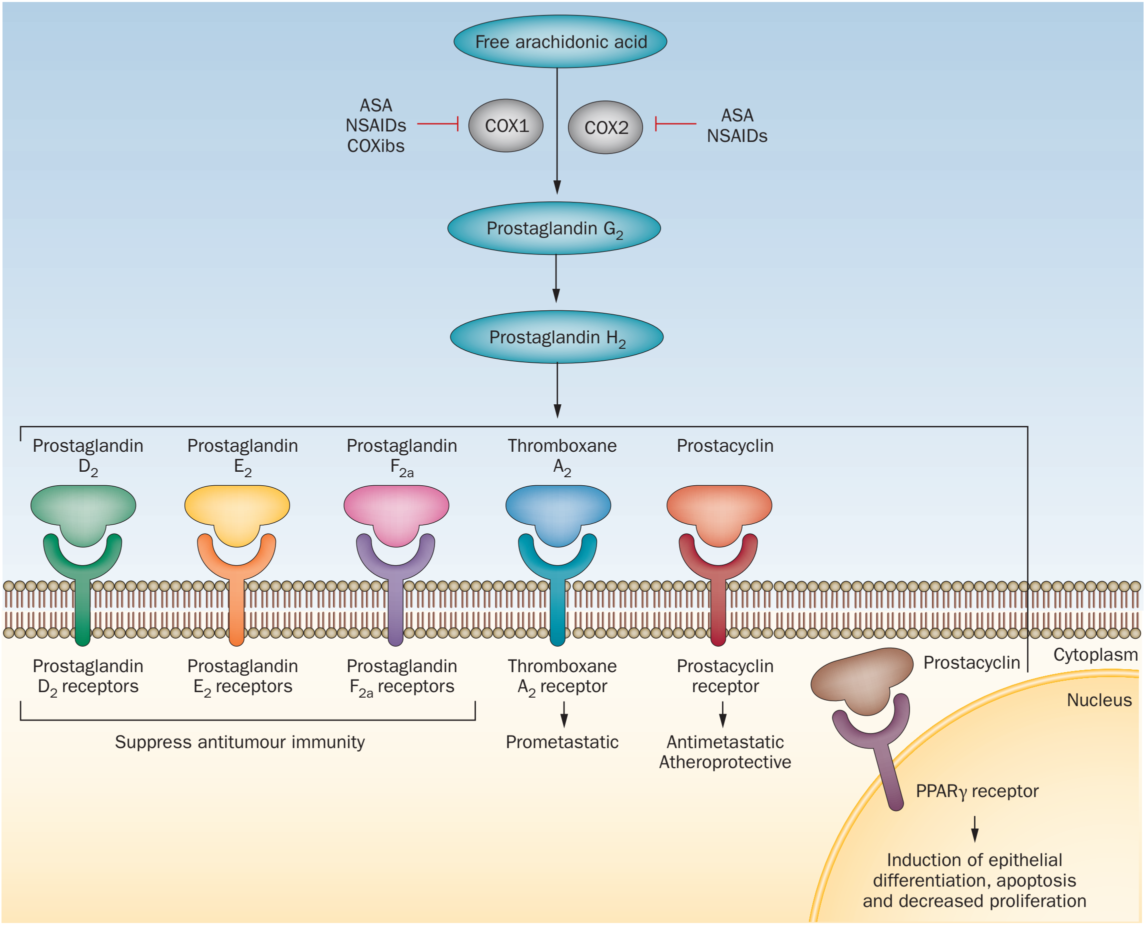


Figure 2 | Schematic pathway of the cyclooxygenase pathway showing conversion of arachidonic acid to prostanoids. The fatty acid arachidonic acid is released from the membrane phospholipids by several forms of phospholipase A_2, which have previously been activated by one of a range of stimuli. The free arachidonic acid is converted to the cyclic endoperoxides prostaglandin G_2 and prostaglandin H_2 by the sequential COX and HOX actions of PGHS-1 or PGHS-2; these isoforms both have dual COX and HOX activity. Aspirin inhibits the conversion of free arachidonic acid to prostaglandin G_2 by inhibiting the COX activity of PGHS-1 or PGHS-2. Prostaglandin H_2 is converted into a range of prostanoids by tissue-specific isomerases; therefore, the inhibition of this pathway prevents (or reduces) the downstream activation of a superfamily of G-protein-coupled receptors by these prostanoids. Prostacyclin binds to both a transmembrane G-protein coupled receptor and to PPARγ, a nuclear receptor. Abbreviations: ASA, acetylsalicylic acid; COX, cyclooxygenase; HOX, hydroperoxidase; NSAIDs, nonsteroidal anti-inflammatory drugs; PGHS, prostaglandin H synthase; PPARγ, peroxisome proliferator-activated receptor γ.

studies and nine prospective cohort studies) also failed to demonstrate a protective effect of aspirin for the prevention of lung cancer.[90] A third meta-analysis of randomized controlled studies of daily aspirin versus no aspirin for the prevention of vascular events in which individual patient data on risk of cancer death was available yielded interesting results.[91] The 20-year risk of death from lung cancer was significantly decreased in trials with scheduled aspirin treatment for 5 years or longer, with hazard ratios (HR) of around 0.70. This effect was entirely due to a reduction in deaths from adenocarcinoma and was seen only in longer-duration studies (5 years or more) with long-term follow up. Although these studies are not definitive, they do raise the hypothesis that to demonstrate the chemopreventive effect of aspirin on lung cancer, more than 5 years of treatment and long-term follow up might be needed.

Pulmonary inflammation is thought to be one factor leading to lung cancer. Therefore, it is perhaps not surprising that systemic and inhaled treatment with the anti-inflammatory agents corticosteroids have chemopreventive efficacy in murine models.[80] A cohort study of patients being treated in Department of Veterans Affairs outpatient clinics reported that those receiving high-dose inhaled corticosteroids and exhibiting good compliance had a reduced risk of lung cancer (HR = 0.39, 95% CI 0.16–0.96) compared to controls.[92] These results are in agreement with a meta-analysis of interventional studies of inhaled corticosteroids assessing COPD outcomes in which a trend towards protection from lung cancer mortality (HR = 0.47, 95% CI 0.22–1.00) was noted (D. Sin, personal communication). The short period of observation of these trials (mean 26 months) limits their applicability to lung cancer chemoprevention.

The thiazolidinedione class of anti-diabetic agents, including rosiglitazone and pioglitazone, act through the PPARγ nuclear receptor, as does the prostacyclin analogue iloprost.[74] Two large observational studies

have compared lung cancer incidence in patients with type II diabetes mellitus treated with thiazolidinediones versus other agents, and both have reported an approximate 33% reduction in lung cancer risk ($P = 0.003$–0.001) in the thiazolidinedione-treated patients.[93,94] However, a third large cohort study from the Kaiser Permanente Northern California Diabetes Registry did not find an association between the use of pioglitazone and decreased risk of lung cancer.[95]

There has been considerable interest in the anti-diabetes agent metformin as a cancer chemoprevention agent. One of the previously described analyses also evaluated metformin use and found a protective effect similar (20–50%, $P = 0.001$) to that observed with the thiazolidinediones.[94] However, published observational studies examining the association between metformin use and lung cancer risk have had mixed results, with two studies reporting that metformin does not have a protective effect.[96–98] By contrast, two systematic reviews and meta-analyses have concluded that the use of metformin in patients with type II diabetes mellitus is associated with significantly lower risk of lung cancer incidence.[99,100] Metformin has also been demonstrated to be chemopreventive in preclinical models.[101] Putative mechanisms for a chemopreventive effect include decreased levels of insulin and insulin-like growth factor and energy stress leading to inhibition of liver kinase B1 (LKB1)/AMP activated protein kinase (AMPK) signalling.[102]

The role of oestrogen in lung carcinogenesis has attracted considerable attention.[103] In the Women's Health Initiative randomized controlled trial, use of oestrogen plus progestin resulted in increased lung cancer mortality.[104] However, use of oestrogen alone in the same study did not increase either incidence or death from lung cancer.[105] An observational study of the California Teachers Study cohort found no association between either oestrogen or oestrogen plus progestin and the risk of lung cancer.[106] A second study of the NIH-AARP Diet and Health Study cohort also found a similar lack of association between oestrogen or oestrogen plus progestin and lung cancer risk.[107] The aromatase inhibitor exemestane was compared to the anti-oestrogen tamoxifen in a randomized controlled trial of women with breast cancer who had been treated for 2–3 years with tamoxifen.[108,109] Exemestane treatment resulted in significantly improved breast cancer disease-free survival and there were fewer lung cancer deaths (4 versus 12) in the exemestane group compared to the tamoxifen group. However, this did not achieve statistical significance owing to the small number of events. Thus, although there is some tantalizing data from both preclinical and clinical studies that treatment with aromatase inhibitors might prevent lung cancer, further investigation is needed.

Finally, a large number of epidemiological studies assessing the relationship between diet and lung cancer incidence have been performed. Almost universally, these demonstrate an inverse correlation between diets that are high in fruit and vegetables and lung cancer incidence.[110] Whether this association represents a true protective action or is the result of other variables associated with fruit and vegetable intake is unclear because no interventional studies have been reported in which dietary manipulations have reduced the incidence of lung cancer.

Future chemoprevention efforts

Agents that have now been clearly shown to lack promise in lung cancer chemoprevention trials, based on null or harmful results in phase III trials, include beta-carotene, aspirin, retinol, multivitamin and mineral supplements, vitamin E, retinyl palmitate, 13-*cis*-retinoic acid, N-acetyl cysteine and selenium (Table 1). Several of these agents are dietary antioxidants, making the general approach of antioxidant administration unattractive without further data supporting efficacy. Aspirin has had negative results in phase III trials, but a recent meta-analysis of randomized trials of aspirin for prevention of vascular events suggests that an intervention period of 5 years or more and prolonged follow up might be necessary to demonstrate a reduction in lung cancer risk.[91]

Chemoprevention agents currently under investigation that have substantial support for incorporation into phase III trials are limited. Inhaled corticosteroids might be the lead agent class because their use is supported by positive data from observational and preclinical studies. However, phase II trials assessing these agents with end points of either dysplasia or nodule growth have been negative,[35] although the lack of specificity surrounding the use of pulmonary nodules as an intermediate end point softens this negative factor. Inhaled corticosteroids are currently used as a treatment for patients with COPD, which is an independent risk factor for lung cancer.[111] Phase III trials with COPD-specific end points have not shown a robust reduction in lung cancer incidence, but the design of these trials, with an average follow-up period of only 26 months, is not ideal for the assessment of cancer chemoprevention.[112–114] Consideration should be given to the initiation of a phase III chemoprevention trial of inhaled corticosteroids for the prevention of lung cancer with a longer duration.

The use of the antidiabetic agents pioglitazone and metformin for chemoprevention is also supported by both preclinical and observational studies (albeit with mixed results). A phase II trial of pioglitazone or placebo with the end point of dysplasia is currently underway.[115] Prostacyclin analogues, specifically iloprost, act through the same PPARγ nuclear receptor as pioglitazone and other thiazolidinediones,[116] so there is some observational support for these agents. In addition, iloprost is effective in preclinical experiments and is the only agent to have demonstrated improvement in endobronchial dysplasia in a phase II trial. The next step to assess the effectiveness of iloprost for chemoprevention would be longer term trials with lung cancer as an end point.

Improved methods for risk assessment, beyond the current models based on clinical and demographic characteristics, would allow chemoprevention trials to be more efficiently designed.[2,3,9] While the current models may be improved by the incorporation of genetics or biomarkers, specific markers have not been validated as providing a major improvement in risk assessment.

Furthermore, the currently available risk models allow the identification of individuals with approximately a 2% yearly risk of developing lung cancer. Currently, the field of lung cancer chemoprevention is more in need of the development of new agents than improved risk models.

Clearly, the translation of preclinical findings into clinically effective chemoprevention of lung cancer has been difficult and largely disappointing. High-throughput screening and functional genomics[117] could be used to identify new promising approaches for chemoprevention, but we are not aware of such applications to date. As lung carcinogenesis is a complex, multistep process involving both the epithelium and stroma, new approaches for lung cancer chemoprevention will still need to be evaluated in preclinical animal models and ultimately in the clinical setting.

The lung is a unique organ, in that there is a long history of therapeutic agents being administered via the inhalational route. Inhalational administration can be more effective than systemic administration and minimizes adverse effects.[118] The potential use of inhaled agents in early phase chemoprevention trials does raise some issues regarding the necessity for and interpretation of traditional end points, such as blood levels of drug. With the exception of inhaled corticosteroids, we are not aware of clinical trials using inhaled agents for lung cancer chemoprevention.[34–36] Furthermore, inhaled treatment in preclinical models is difficult and often results in different patterns of drug distribution compared to those seen in humans.[119] However, inhaled chemoprevention is an attractive approach going forward as it maximizes drug delivery to the target organ and minimizes systemic effects.[118]

Personalized medicine has been exemplified by the success of targeted therapy in treating molecularly defined subsets of patients, most commonly with specific driver mutations. The complexity of genetic mutations in both adenocarcinoma and squamous-cell lung cancer,[32,33] coupled with the difficulty in detecting and analysing premalignant lesions, are challenges to personalized chemoprevention for lung cancer. It might be possible to define common pathways in early carcinogenesis, such as PI3K activation or p53 inactivation, which are shared by significant fractions of the at-risk population and might be targeted in a directed fashion provided effective agents are identified. Myoinositol has been proposed as an agent that may be useful for targeting individuals specifically with premalignant lesions characterized by activated PI3K signalling, and an ongoing trial is testing this hypothesis.[61] Alternately, phenotypes such as airway or parenchymal inflammation, incipient angiogenesis, tissue hypoxia, or excessive growth factor expression, might be defined, each of which would respond best to a different intervention.[82,120,121]

Conclusions

Lung cancer is the leading cause of death from cancer worldwide and a disease for which high-risk individuals can be readily identified. Even after smoking cessation is accomplished, former smokers are at significant residual risk of developing lung cancer. These factors make lung cancer an ideal target for chemoprevention. Attempts to discover and validate chemopreventive agents have been frustrating; however, there are a number of interesting compounds that might be effective. These include inhaled glucocorticoids, myoinositol, prostacyclin analogues and thiazolidinediones. Advances in driver-mutation targeted treatment of lung cancer might provide insight into targets and potential chemotherapeutic agents for future trials. Personalized chemoprevention is likely to be targeted more to phenotypes, such as airway inflammation, alveolar hypoxia, or incipient angiogenesis, than to specific driver mutations. Validating biomarkers of risk and response will be critical to advancing the field.

Review criteria

The PubMed database was searched for articles with no restriction on date. The search terms used included: “lung cancer”, “chemoprevention”, “animal models”, and the various chemopreventive agents in past and current trials (summarized in Tables 1 and 2), “trials”, “biomarkers”, “tobacco smoke”, “pre-malignant airway lesions” and “carcinogenesis”. Only articles published in English were considered, and the final PubMed search was conducted on 2 April 2013.

1. Sporn, M. B. Approaches to prevention of epithelial cancer during the preneoplastic period. *Cancer Res.* **36**, 2699–2702 (1976).
2. Spitz, M. R. *et al.* A risk model for prediction of lung cancer. *J. Natl Cancer Inst.* **99**, 715–726 (2007).
3. Bach, P. B. *et al.* Variations in lung cancer risk among smokers. *J. Natl Cancer Inst.* **95**, 470–478 (2003).
4. Anthonisen, N. R. *et al.* The effects of a smoking cessation intervention on 14.5-year mortality: a randomized clinical trial. *Ann. Intern. Med.* **142**, 233–239 (2005).
5. Jemal, A. *et al.* Global cancer statistics. *CA Cancer J. Clin.* **61**, 69–90 (2011).
6. Siegel, R., Naishadham, D. & Jemal, A. Cancer statistics, 2012. *CA Cancer J. Clin.* **62**, 10–29 (2012).
7. National Lung Screening Trial Research Team *et al.* Reduced lung-cancer mortality with low-dose computed tomographic screening. *N. Engl. J. Med.* **365**, 395–409 (2011).
8. Johnson, B. E. Second lung cancers in patients after treatment for an initial lung cancer. *J. Natl Cancer Inst.* **90**, 1335–1345 (1998).
9. Spitz, M. R. *et al.* An expanded risk prediction model for lung cancer. *Cancer Prev. Res. (Phila.)* **1**, 250–254 (2008).
10. The Health Consequences of Smoking: A Report of the Surgeon General. Office of the Surgeon General (US); Office on Smoking and Health (US). Atlanta (GA): Centers for Disease Control and Prevention (US) [online], http://www.surgeongeneral.gov/library/reports/smokingconsequences/index.html (2004).
11. Jha, P. *et al.* 21st-century hazards of smoking and benefits of cessation in the United States. *N. Engl. J. Med.* **368**, 341–350 (2013).
12. Pirie, K. *et al.* The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. *Lancet* **381**, 133–141 (2012).
13. Tong, L., Spitz, M. R., Fueger, J. J. & Amos, C. A. Lung carcinoma in former smokers. *Cancer* **78**, 1004–1010 (1996).
14. Halpern, M. T., Gillespie, B. W. & Warner, K. E. Patterns of absolute risk of lung cancer mortality in former smokers. *J. Natl Cancer Inst.* **85**, 457–464 (1993).
15. Hennekens, C. H. *et al.* Lack of effect of long-term supplementation with beta carotene on the incidence of malignant neoplasms and cardiovascular disease. *N. Engl. J. Med.* **334**, 1145–1149 (1996).
16. Omenn, G. S. *et al.* Risk factors for lung cancer and for intervention effects in CARET, the Beta-Carotene and Retinol Efficacy Trial. *J. Natl Cancer Inst.* **88**, 1550–1559 (1996).

17. Lockwood, W. W. *et al.* Divergent genomic and epigenomic landscapes of lung cancer subtypes underscore the selection of different oncogenic pathways during tumor development. *PLoS ONE* **7**, e37775 (2012).
18. Nicholson, A. G. *et al.* Reproducibility of the WHO/IASLC grading system for pre-invasive squamous lesions of the bronchus: a study of inter-observer and intra-observer variation. *Histopathology* **38**, 202–208 (2001).
19. Jonsson, S. *et al.* Chromosomal aneusomy in bronchial high-grade lesions is associated with invasive lung cancer. *Am. J. Respir. Crit. Care Med.* **177**, 342–347 (2008).
20. Salaun, M. *et al.* Molecular predictive factors for progression of high-grade preinvasive bronchial lesions. *Am. J. Respir. Crit. Care Med.* **177**, 880–886 (2008).
21. van Boerdonk, R. A. *et al.* DNA copy number alterations in endobronchial squamous metaplastic lesions predict lung cancer. *Am. J. Respir. Crit. Care Med.* **184**, 948–956 (2011).
22. Franklin, W. A. *et al.* Widely dispersed p53 mutation in respiratory epithelium. A novel mechanism for field carcinogenesis. *J. Clin. Invest.* **100**, 2133–2137 (1997).
23. Wistuba, I. I. *et al.* Sequential molecular abnormalities are involved in the multistage development of squamous cell lung carcinoma. *Oncogene* **18**, 643–650 (1999).
24. McCaughan, F. *et al.* Progressive 3q amplification consistently targets SOX2 in preinvasive squamous lung cancer. *Am. J. Respir. Crit. Care Med.* **182**, 83–91 (2010).
25. Travis, W. D. *et al.* Bronchioloalveolar carcinoma and lung adenocarcinoma: the clinical importance and research relevance of the 2004 World Health Organization pathologic criteria. *J. Thorac. Oncol.* **1**, S13–S19 (2006).
26. Tang, X. *et al.* EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. *Cancer Res.* **65**, 7568–7572 (2005).
27. Meert, A. P. *et al.* Epidermal growth factor receptor expression in pre-invasive and early invasive bronchial lesions. *Eur. Respir. J.* **21**, 611–615 (2003).
28. Soh, J. *et al.* Sequential molecular changes during multistage pathogenesis of small peripheral adenocarcinomas of the lung. *J. Thorac. Oncol.* **3**, 340–347 (2008).
29. Slaughter, D. P., Southwick, H. W. & Smejkal, W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* **6**, 963–968 (1953).
30. Rusch, V., Klimstra, D., Linkov, I. & Dmitrovsky, E. Aberrant expression of p53 or the epidermal growth factor receptor is frequent in early bronchial neoplasia and coexpression precedes squamous cell carcinoma development. *Cancer Res.* **55**, 1365–1372 (1995).
31. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
32. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525 (2012).
33. Imielinski, M. *et al.* Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107–1120 (2012).
34. Veronesi, G. *et al.* Randomized phase II trial of inhaled budesonide versus placebo in high-risk individuals with CT screen-detected lung nodules. *Cancer Prev. Res. (Phila.)* **4**, 34–42 (2011).
35. Lam, S. *et al.* A randomized phase IIb trial of pulmicort turbuhaler (budesonide) in people with dysplasia of the bronchial epithelium. *Clin. Cancer Res.* **10**, 6502–6511 (2004).
36. van den Berg, R. M. *et al.* CT detected indeterminate pulmonary nodules in a chemoprevention trial of fluticasone. *Lung Cancer* **60**, 57–61 (2008).
37. Final report on the aspirin component of the ongoing Physicians' Health Study. Steering Committee of the Physicians' Health Study Research Group. *N. Engl. J. Med.* **32**, 129–135 (1989).
38. Cook, N. R. *et al.* Low-dose aspirin in the primary prevention of cancer: the Women's Health Study: a randomized controlled trial. *JAMA* **294**, 47–55 (2005).
39. Peto, R. *et al.* Randomised trial of prophylactic daily aspirin in British male doctors. *Br. Med. J. (Clin. Res. Ed.)* **296**, 313–316 (1988).
40. van Zandwijk, N., Dalesio, O., Pastorino, U., de Vries, N. & van Tinteren, H. EUROSCAN, a randomized trial of vitamin A and N-acetylcysteine in patients with head and neck cancer or lung cancer. For the European Organization for Research and Treatment of Cancer Head and Neck and Lung Cancer Cooperative Groups. *J. Natl Cancer Inst.* **92**, 977–986 (2000).
41. Lippman, S. M. *et al.* Randomized phase III intergroup trial of isotretinoin to prevent second primary tumors in stage I non-small-cell lung cancer. *J. Natl Cancer Inst.* **93**, 605–618 (2001).
42. Blumberg, J. & Block, G. The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study in Finland. *Nutr. Rev.* **52**, 242–245 (1994).
43. Slatore, C. G., Littman, A. J., Au, D. H., Satia, J. A. & White, E. Long-term use of supplemental multivitamins, vitamin C, vitamin E, and folate does not reduce the risk of lung cancer. *Am. J. Respir. Crit. Care Med.* **177**, 524–530 (2008).
44. Karp, D. D. *et al*. A phase III, intergroup, randomized, double-blind, chemoprevention trial of selenium (Se) supplementation in resected stage I non-small cell lung cancer (NSCLC) [abstract]. *J. Clin. Oncol.* **18** (Suppl.), CRA7004 (2010).
45. Kelly, R. J., Lopez-Chavez, A. & Szabo, E. Criteria of evidence to move potential chemopreventive agents into late phase clinical trials. *Curr. Drug Targets* **12**, 1983–1988 (2011).
46. Keith, R. L. & Miller, Y. E. Lung cancer: genetics of risk and advances in chemoprevention. *Curr. Opin. Pulm. Med.* **11**, 265–271 (2005).
47. Prentice, R. L. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat. Med.* **8**, 431–440 (1989).
48. Miller, Y. E. *et al.* Bronchial epithelial Ki-67 index is related to histology, smoking, and gender, but not lung cancer or chronic obstructive pulmonary disease. *Cancer Epidemiol. Biomarkers Prev.* **16**, 2425–2431 (2007).
49. Spira, A. *et al.* Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proc. Natl Acad. Sci. USA* **101**, 10143–10148 (2004).
50. Spira, A. *et al.* Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat. Med.* **13**, 361–366 (2007).
51. Rahman, S. M. *et al.* Lung cancer diagnosis from proteomic analysis of preinvasive lesions. *Cancer Res.* **71**, 3009–3017 (2011).
52. Ostroff, R. M. *et al.* Unlocking biomarker discovery: large scale application of aptamer proteomic technology for early detection of lung cancer. *PLoS ONE* **5**, e15003 (2010).
53. Peled, N. *et al.* Non-invasive breath analysis of pulmonary nodules. *J. Thorac. Oncol.* **7**, 1528–1533 (2012).
54. Kelly, K. *et al.* A randomized phase II chemoprevention trial of 13-CIS retinoic acid with or without alpha tocopherol or observation in subjects at high risk for lung cancer. *Cancer Prev. Res. (Phila.)* **2**, 440–449 (2009).
55. Lee, J. S. *et al.* Randomized placebo-controlled trial of isotretinoin in chemoprevention of bronchial squamous metaplasia. *J. Clin. Oncol.* **12**, 937–945 (1994).
56. Keith, R. L. *et al.* Oral iloprost improves endobronchial dysplasia in former smokers. *Cancer Prev. Res. (Phila.)* **4**, 793–802 (2011).
57. Kim, E. S. *et al.* Biological activity of celecoxib in the bronchial epithelium of current and former smokers. *Cancer Prev. Res. (Phila.)* **3**, 148–159 (2010).
58. Mao, J. T. *et al.* Lung cancer chemoprevention with celecoxib in former smokers. *Cancer Prev. Res. (Phila.)* **4**, 984–993 (2011).
59. Chang, B. *et al.* Natural history of pure ground-glass opacity lung nodules detected by low-dose CT scan. *Chest* **143**, 172–178 (2013).
60. Lam, S. *et al.* A phase I study of myo-inositol for lung cancer chemoprevention. *Cancer Epidemiol. Biomarkers Prev.* **15**, 1526–1531 (2006).
61. Gustafson, A. M. *et al.* Airway PI3K pathway activation is an early and reversible event in lung cancer development. *Sci. Transl. Med.* **2**, 26ra25 (2010).
62. Malkinson, A. M., Koski, K. M., Evans, W. A. & Festing, M. F. Butylated hydroxytoluene exposure is necessary to induce lung tumors in BALB mice treated with 3-methylcholanthrene. *Cancer Res.* **57**, 2832–2824 (1997).
63. Johnson, L. *et al.* Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature* **410**, 1111–1116 (2001).
64. Regales, L. *et al.* Development of new mouse lung tumor models expressing EGFR T790M mutants associated with clinical resistance to kinase inhibitors. *PLoS ONE* **2**, e810 (2007).
65. Fisher, G. H. *et al.* Induction and apoptotic regression of lung adenocarcinomas by regulation of a K-Ras transgene in the presence and absence of tumor suppressor genes. *Genes Dev.* **15**, 3249–3262 (2001).
66. Jackson, E. L. *et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev.* **15**, 3243–3248 (2001).
67. Mirvish, S. S. The carcinogenic action and metabolism of urethan and N-hydroxyurethan. *Adv. Cancer Res.* **1**, 1–42 (1968).
68. Witschi, H. Tobacco smoke as a mouse lung carcinogen. *Exp. Lung Res.* **24**, 385–394 (1998).
69. Wang, Y. *et al.* A chemically induced model for squamous cell carcinoma of the lung in mice: histopathology and strain susceptibility. *Cancer Res.* **64**, 1647–1654 (2004).
70. Hudish, T. M. *et al.* N-nitroso-tris-chloroethylurea induces premalignant squamous dysplasia in mice. *Cancer Prev. Res. (Phila.)* **5**, 283–289 (2012).
71. Stearman, R. S. *et al.* Analysis of orthologous gene expression between human pulmonary adenocarcinoma and a carcinogen-induced murine model. *Am. J. Pathol.* **167**, 1763–1775 (2005).
72. Wang, D. & DuBois, R. N. Eicosanoids and cancer. *Nat. Rev. Cancer* **10**, 181–193 (2010).
73. Vane, J. R. Prostacyclin: a hormone with a therapeutic potential. The Sir Henry Dale Lecture for 1981. *J. Endocrinol.* **95**, 3P–43P (1982).
74. Nemenoff, R. *et al.* Prostacyclin prevents murine lung cancer independent of the membrane receptor by activation of peroxisomal proliferator–activated receptor gamma. *Cancer Prev. Res. (Phila.)* **1**, 349–356 (2008).

75. Chang, T. H. & Szabo, E. Induction of differentiation and apoptosis by ligands of peroxisome proliferator-activated receptor gamma in non-small cell lung cancer. *Cancer Res.* **60**, 1129–1138 (2000).
76. Keshamouni, V. G. *et al.* Peroxisome proliferator-activated receptor-gamma activation inhibits tumor progression in non-small-cell lung cancer. *Oncogene* **23**, 100–108 (2004).
77. Lyon, C. M. *et al.* Rosiglitazone prevents the progression of preinvasive lung cancer in a murine model. *Carcinogenesis* **30**, 2095–2099 (2009).
78. Wang, Y. *et al.* Chemopreventive effects of pioglitazone on chemically induced lung carcinogenesis in mice. *Mol. Cancer Ther.* **9**, 3074–3082 (2010).
79. Fu, H. *et al.* Chemoprevention of lung carcinogenesis by the combination of aerosolized budesonide and oral pioglitazone in A/J mice. *Mol. Carcinog.* **50**, 913–921 (2011).
80. Wattenberg, L. W. *et al.* Chemoprevention of pulmonary carcinogenesis by brief exposures to aerosolized budesonide or beclomethasone dipropionate and by the combination of aerosolized budesonide and dietary myo-inositol. *Carcinogenesis* **21**, 179–182 (2000).
81. Keith, R. L. *et al.* Manipulation of pulmonary prostacyclin synthase expression prevents murine lung cancer. *Cancer Res.* **62**, 734–740 (2002).
82. Karoor, V., Le, M., Merrick, D., Dempsey, E. C. & Miller, Y. E. Vascular endothelial growth factor receptor 2-targeted chemoprevention of murine lung tumors. *Cancer Prev. Res. (Phila.)* **3**, 1141–1147 (2010).
83. Stabile, L. P. *et al.* Prevention of tobacco carcinogen-induced lung cancer in female mice using antiestrogens. *Carcinogenesis* **33**, 2181–2189 (2012).
84. Moody, T. W. *et al.* Indomethacin reduces lung adenoma number in A/J mice. *Anticancer Res.* **21**, 1749–1755 (2001).
85. Kisley, L. R. *et al.* Celecoxib reduces pulmonary inflammation but not lung tumorigenesis in mice. *Carcinogenesis* **23**, 1653–1660 (2002).
86. Liby, K. T. & Sporn, M. B. Synthetic oleanane triterpenoids: multifunctional drugs with a broad range of applications for prevention and treatment of chronic disease. *Pharmacol. Rev.* **64**, 972–1003 (2012).
87. Liby, K. *et al.* The rexinoid LG100268 and the synthetic triterpenoid CDDO-methyl amide are more potent than erlotinib for prevention of mouse lung carcinogenesis. *Mol. Cancer Ther.* **7**, 1251–1257 (2008).
88. Liby, K. *et al.* The synthetic triterpenoids CDDO-methyl ester and CDDO-ethyl amide prevent lung cancer induced by vinyl carbamate in A/J mice. *Cancer Res.* **67**, 2414–2419 (2007).
89. Algra, A. M. & Rothwell, P. M. Effects of regular aspirin on long-term cancer incidence and metastasis: a systematic comparison of evidence from observational studies versus randomised trials. *Lancet Oncol.* **13**, 518–527 (2012).
90. Oh, S. W. *et al.* Aspirin use and risk for lung cancer: a meta-analysis. *Ann. Oncol.* **22**, 2456–2465 (2011).
91. Rothwell, P. M. *et al.* Effect of daily aspirin on long-term risk of death due to cancer: analysis of individual patient data from randomised trials. *Lancet* **377**, 31–41 (2011).
92. Parimon, T. *et al.* Inhaled corticosteroids and risk of lung cancer among patients with chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* **175**, 712–719 (2007).
93. Govindarajan, R. *et al.* Thiazolidinediones and the risk of lung, prostate, and colon cancer in patients with diabetes. *J. Clin. Oncol.* **25**, 1476–1481 (2007).
94. Mazzone, P. J. *et al.* The effect of metformin and thiazolidinedione use on lung cancer in diabetics. *BMC Cancer* **12**, 410 (2012).
95. Ferrara, A. *et al.* Cohort study of pioglitazone and cancer incidence in patients with diabetes. *Diabetes Care* **34**, 923–929 (2011).
96. Smiechowski, B. B., Azoulay, L., Yin, H., Pollak, M. N. & Suissa, S. The use of metformin and the incidence of lung cancer in patients with type 2 diabetes. *Diabetes Care* **36**, 124–129 (2013).
97. Bodmer, M., Becker, C., Jick, S. S. & Meier, C. R. Metformin does not alter the risk of lung cancer: a case-control analysis. *Lung Cancer* **78**, 133–137 (2012).
98. Lai, S. W. *et al.* Antidiabetes drugs correlate with decreased risk of lung cancer: a population-based observation in Taiwan. *Clin. Lung Cancer* **13**, 143–148 (2012).
99. Noto, H., Goto, A., Tsujimoto, T. & Noda, M. Cancer risk in diabetic patients treated with metformin: a systematic review and meta-analysis. *PLoS ONE* **7**, e33411 (2012).
100. Soranna, D. *et al.* Cancer risk associated with use of metformin and sulfonylurea in type 2 diabetes: a meta-analysis. *Oncologist* **17**, 813–822 (2012).
101. Memmott, R. M. *et al.* Metformin prevents tobacco carcinogen-induced lung tumorigenesis. *Cancer Prev. Res. (Phila.)* **3**, 1066–1076 (2010).
102. Engelman, J. A. & Cantley, L. C. Chemoprevention meets glucose control. *Cancer Prev. Res. (Phila.)* **3**, 1049–1052 (2010).
103. Siegfried, J. M., Hershberger, P. A. & Stabile, L. P. Estrogen receptor signaling in lung cancer. *Semin. Oncol.* **36**, 524–531 (2009).
104. Ganti, A. K., Sahmoun, A. E., Panwalkar, A. W., Tendulkar, K. K. & Potti, A. Hormone replacement therapy is associated with decreased survival in women with lung cancer. *J. Clin. Oncol.* **24**, 59–63 (2006).
105. Chlebowski, R. T. *et al.* Lung cancer among postmenopausal women treated with estrogen alone in the women's health initiative randomized trial. *J. Natl Cancer Inst.* **102**, 1413–1421 (2010).
106. Clague, J. *et al.* Menopausal hormone therapy does not influence lung cancer risk: results from the California Teachers Study. *Cancer Epidemiol. Biomarkers Prev.* **20**, 560–564 (2011).
107. Brinton, L. A. *et al.* Unopposed estrogen and estrogen plus progestin menopausal hormone therapy and lung cancer risk in the NIH-AARP Diet and Health Study Cohort. *Cancer Causes Control* **23**, 487–496 (2012).
108. Coombes, R. C. *et al.* A randomized trial of exemestane after two to three years of tamoxifen therapy in postmenopausal women with primary breast cancer. *N. Engl. J. Med.* **350**, 1081–1092 (2004).
109. Goss, P. E. *et al.* Exemestane for breast-cancer prevention in postmenopausal women. *N. Engl. J. Med.* **364**, 2381–2391 (2011).
110. Key, T. J. Fruit and vegetables and cancer risk. *Br. J. Cancer* **104**, 6–11 (2011).
111. Young, R. P. *et al.* COPD prevalence is increased in lung cancer, independent of age, sex and smoking history. *Eur. Respir. J.* **34**, 380–386 (2009).
112. Burge, P. S. *et al*. Randomised, double blind, placebo controlled study of fluticasone propionate in patients with moderate to severe chronic obstructive pulmonary disease: the ISOLDE trial. *BMJ* **320**, 1297–1303 (2000).
113. Lung Health Study Research Group. Effect of inhaled triamcinolone on the decline in pulmonary function in chronic obstructive pulmonary disease. *N. Engl. J. Med*. **343**, 1902–1909 (2000).
114. Vestbo, J. *et al*. Long-term effect of inhaled budesonide in mild and moderate chronic obstructive pulmonary disease: a randomised controlled trial. *Lancet* **353**, 1819–1823 (1999).
115. US National Library of Medicine. *ClinicalTrials.gov* [online], http://clinicaltrials.gov/show/NCT00780234 (2013).
116. Gupta, R. A. *et al.* Prostacyclin-mediated activation of peroxisome proliferator-activated receptor delta in colorectal cancer. *Proc. Natl Acad. Sci. USA* **97**, 13275–13280 (2000).
117. Alvarez-Calderon, F., Gregory, M. A. & Degregori, J. Using functional genomics to overcome therapeutic resistance in hematological malignancies. *Immunol. Res.* **55**, 100–115 (2013).
118. Wang, D. L. *et al.* Topical delivery of 13-cis-retinoic acid by inhalation up-regulates expression of rodent lung but not liver retinoic acid receptors. *Clin. Cancer Res.* **6**, 3636–3645 (2000).
119. Mendez, L. B., Gookin, G. & Phalen, R. F. Inhaled aerosol particle dosimetry in mice: a review. *Inhal. Toxicol*. **22**, 1032–1037 (2010).
120. Keith, R. L. *et al*. Angiogenic squamous dysplasia in bronchi of individuals at high risk for lung cancer. *Clin. Cancer Res*. **6**, 1616–1625 (2000).
121. Karoor, V. *et al*. Alveolar hypoxia promotes murine lung tumor growth through a VEGFR-2/EGFR-dependent mechanism. *Cancer Prev. Res. (Phila.)* **5**, 1061–1071 (2012).
122. Omenn, G. S. *et al.* Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. *N. Engl. J. Med.* **334**, 1150–1155 (1996).
123. Kurie, J. M. *et al.* N-(4-hydroxyphenyl)retinamide in the chemoprevention of squamous metaplasia and dysplasia of the bronchial epithelium. *Clin. Cancer Res.* **6**, 2973–2979 (2000).
124. Arnold, A. M. *et al.* The effect of the synthetic retinoid etretinate on sputum cytology: results from a randomised trial. *Br. J. Cancer* **65**, 737–743 (1992).
125. McLarty, J. W. *et al.* Beta-carotene, vitamin A, and lung cancer chemoprevention: results of an intermediate endpoint study. *Am. J. Clin. Nutr.* **62**, 1431S–1438S (1995).
126. Heimburger, D. C. *et al.* Improvement in bronchial squamous metaplasia in smokers treated with folate and vitamin B12. Report of a preliminary randomized, double-blind intervention trial. *JAMA* **259**, 1525–3150 (1988).
127. Lam, S. *et al.* A randomized phase IIb trial of anethole dithiolethione in smokers with bronchial dysplasia. *J. Natl Cancer Inst.* **94**, 1001–1009 (2002).

Acknowledgements
The authors acknowledge support from the Department of Veterans Affairs Merit Review Program; NCI (Colorado SPORE in Lung Cancer—P50 CA58187 and R01 CA164780-02); and an Early Detection Award from the LUNGevity Foundation.

Author contributions
Both authors researched data for the article, made a substantial contribution to the discussion of the content, wrote the article and edited it prior to submission.

Rare variants of large effect in *BRCA2* and *CHEK2* affect risk of lung cancer

Yufei Wang[1,61], James D McKay[2,61,62], Thorunn Rafnar[3], Zhaoming Wang[4], Maria N Timofeeva[2], Peter Broderick[1], Xuchen Zong[5], Marina Laplana[6], Yongyue Wei[7], Younghun Han[8], Amy Lloyd[1], Manon Delahaye-Sourdeix[2], Daniel Chubb[1], Valerie Gaborieau[2], William Wheeler[9], Nilanjan Chatterjee[4], Gudmar Thorleifsson[3], Patrick Sulem[3], Geoffrey Liu[10], Rudolf Kaaks[11,12], Marc Henrion[1], Ben Kinnersley[1], Maxime Vallée[2], Florence LeCalvez-Kelm[2], Victoria L Stevens[13], Susan M Gapstur[13], Wei V Chen[14], David Zaridze[15], Neonilia Szeszenia-Dabrowska[16], Jolanta Lissowska[17], Peter Rudnai[18], Eleonora Fabianova[19], Dana Mates[20], Vladimir Bencko[21], Lenka Foretova[22], Vladimir Janout[23], Hans E Krokan[24], Maiken Elvestad Gabrielsen[24], Frank Skorpen[25], Lars Vatten[26], Inger Njølstad[27], Chu Chen[28], Gary Goodman[28], Simone Benhamou[29], Tonu Vooder[30], Kristjan Välk[31], Mari Nelis[32,33], Andres Metspalu[32], Marcin Lener[34], Jan Lubiński[34], Mattias Johansson[2], Paolo Vineis[35,36], Antonio Agudo[37], Francoise Clavel-Chapelon[38–40], H Bas Bueno-de-Mesquita[35,41,42], Dimitrios Trichopoulos[43–45], Kay-Tee Khaw[46], Mikael Johansson[47], Elisabete Weiderpass[48–51], Anne Tjønneland[52], Elio Riboli[35], Mark Lathrop[53], Ghislaine Scelo[2], Demetrius Albanes[4], Neil E Caporaso[4], Yuanqing Ye[54], Jian Gu[54], Xifeng Wu[54], Margaret R Spitz[55], Hendrik Dienemann[12,56], Albert Rosenberger[57], Li Su[7], Athena Matakidou[58], Timothy Eisen[59,60], Kari Stefansson[3], Angela Risch[6,12], Stephen J Chanock[4], David C Christiani[7], Rayjean J Hung[5], Paul Brennan[2], Maria Teresa Landi[4,61,62], Richard S Houlston[1,61,62] & Christopher I Amos[8,61,62]

We conducted imputation to the 1000 Genomes Project of four genome-wide association studies of lung cancer in populations of European ancestry (11,348 cases and 15,861 controls) and genotyped an additional 10,246 cases and 38,295 controls for follow-up. We identified large-effect genome-wide associations for squamous lung cancer with the rare variants BRCA2 p.Lys3326X (rs11571833, odds ratio (OR) = 2.47, $P = 4.74 \times 10^{-20}$) and CHEK2 p.Ile157Thr (rs17879961, OR = 0.38, $P = 1.27 \times 10^{-13}$). We also showed an association between common variation at 3q28 (*TP63*, rs13314271, OR = 1.13, $P = 7.22 \times 10^{-10}$) and lung adenocarcinoma that had been previously reported only in Asians. These findings provide further evidence for inherited genetic susceptibility to lung cancer and its biological basis. Additionally, our analysis demonstrates that imputation can identify rare disease-causing variants with substantive effects on cancer risk from preexisting genome-wide association study data.

Lung cancer causes over 1 million deaths each year worldwide[1]. Although primarily caused by tobacco smoking, studies have also implicated inherited genetic factors in the etiology of lung cancer; notably, genome-wide association studies (GWAS) in Europeans have consistently identified polymorphic variation at 15q25.1 (*CHRNA5-CHRNA3-CHRNB4*), 5p15.33 (*TERT-CLPTM1*) and 6p21.33 (*BAG6* (also called *BAT3*)-*MSH5*) as determinants of lung cancer risk[2–6]. Additionally, susceptibility loci for lung cancer at 3q28, 6q22.2, 13q12.12, 10q25.2 and 22q12.2 in Asians have been identified through GWAS[7–9].

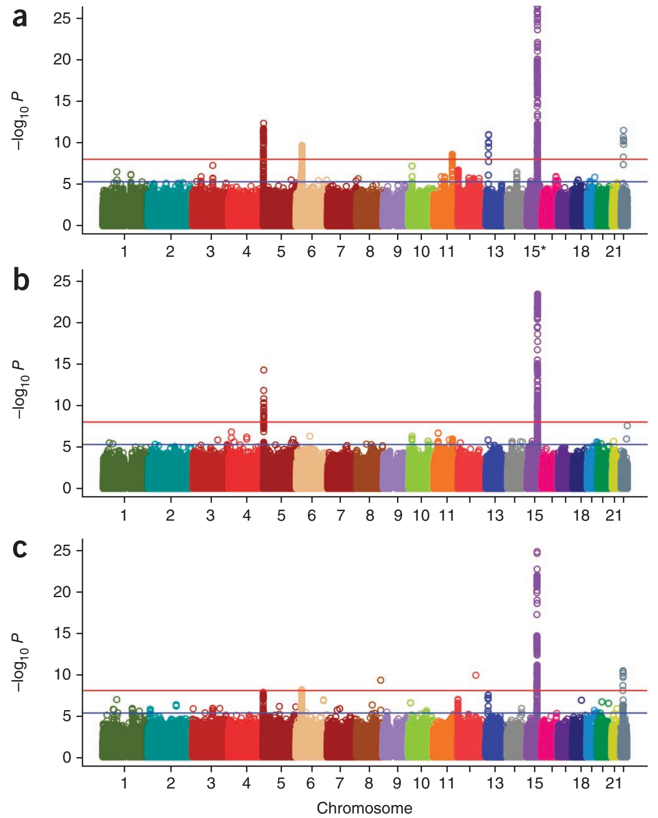
Non–small cell lung cancer (NSCLC) is the most common lung cancer histology, comprised primarily of adenocarcinoma (AD) and squamous cell carcinoma (SQ). These lung cancer histologies have different molecular characteristics that reflect differences in etiology and carcinogenesis[10]. Perhaps not surprisingly, there is variability in the genetic effects on lung cancer risk by histology, with subtype-specific associations at 5p15.33 (*TERT-CLPTM1*) for AD[11,12] and at 9p21 (*CDKN2A/CDKN2B*)[13] and 12q13.33 (*RAD52*)[14] for SQ. In addition, the 6p21.33 associations are stronger for SQ than for AD[13].

To identify additional lung cancer susceptibility loci, we conducted a meta-analysis of four lung cancer GWAS in populations of European ancestry: the MD Anderson Cancer Center (MDACC) GWAS, the Institute of Cancer Research (ICR) GWAS, the National Cancer Institute (NCI) GWAS and the International Agency for Research on Cancer (IARC) GWAS (Online Methods), which were genotyped using Illumina HumanHap 317, 317+240S, 370Duo, 550, 610 or 1M arrays (**Supplementary Table 1**). After filtering, the studies provided genotypes on 11,348 cases and 15,861 controls (**Supplementary Table 1**). Before undertaking meta-analysis of the GWAS data, we searched for potential errors and biases in the data sets. Quantile-quantile (Q-Q) plots of genome-wide association test statistics showed minimal inflation,

A full list of author affiliations appears at the end of the paper.

Received 25 September 2013; accepted 8 May 2014; published online 1 June 2014; doi:10.1038/ng.3002

Figure 1 Genome-wide P values plotted against their respective chromosomal positions. (**a**–**c**) All lung cancer (**a**), AD (**b**) and SQ (**c**). Shown are the genome-wide P values (two sided) obtained using the Cochran-Armitage trend test from analysis of 8.9 million successfully imputed autosomal SNPs in 11,348 cases and 15,861 controls from the discovery phase. The red and blue horizontal lines represent the significance thresholds of $P = 5.0 \times 10^{-8}$ and $P = 5.0 \times 10^{-6}$, respectively. Any region that contained at least one association signal better than $P = 5.0 \times 10^{-6}$ was selected for the *in silico* replication.



rendering substantial cryptic population substructure or differential genotype calling between cases and controls unlikely (λ = 1.01–1.05; **Supplementary Fig. 1**). To bring genotype data obtained from different arrays into a common platform and recover untyped genotypes, we imputed >10 million SNPs using 1000 Genomes Project data as the reference. Q-Q plots for all SNPs and those restricted to rare SNPs (minor allele frequency (MAF) <1%) after imputation did not show evidence of substantive overdispersion introduced by imputation (λ = 0.99–1.06 and λ = 0.82–1.05, respectively; **Supplementary Fig. 1**).

Pooling data from each GWAS, we derived joint ORs and 95% confidence intervals (CIs) under a fixed-effects model for each SNP and the associated per-allele P values. To explore variability in associations according to tumor histology, we derived ORs for all lung cancer, AD and SQ.

Our meta-analysis identified 50 SNPs that showed evidence of association with lung cancer, AD or SQ ($P < 5.0 \times 10^{-6}$; **Fig. 1**) at loci not reported previously in Europeans (**Fig. 1**). We evaluated 1-Mb regions encompassing these 50 SNPs for association through *in silico* replication in the Harvard[15] and deCODE[16] series. Nine of the SNPs within these 50 regions showed support for an association (combined $P < 5.0 \times 10^{-7}$). We attempted genotyping of these nine SNPs in four additional series: the Heidelberg–European Prospective Investigation into Cancer and Nutrition (EPIC), ICR, IARC and Toronto replications (**Supplementary Table 2b** and Online Methods). rs185577307 could not be genotyped because of repetitive sequence. Collectively, genotypes were available from 21,594 cases and 54,156 controls, providing 80% power to detect a variant with MAF of 0.01 and conferring a relative risk of ≥1.5. In the combined analysis of all GWAS plus replication series data, SNPs mapping to 13q13.1 (rs11571833 and rs56084662), 22q12.1 (rs17879961) and 3q28 (rs13314271) showed evidence for association, which was statistically significant after adjustment for multiple testing ($P < 3.0 \times 10^{-9}$; **Fig. 2** and **Supplementary Table 3**). We confirmed the high fidelity of imputation by genotyping rs11571833, rs17879961 and rs13314271 in subsets of the ICR, IARC, NCI and MDACC GWAS (**Supplementary Table 2** and Online Methods). The NCI GWAS comprised samples from Finland, Italy and the United States. The IARC GWAS comprised samples from ten series from western and eastern Europe and the United States. Although adjustment of test statistics for principal components generated on common SNPs had been applied to these GWAS, confounding of rare variants in spatially structured populations is not necessarily corrected by such methods[17]. We therefore investigated whether country of origin had an impact on the associations at 13q13.1 and 22q12.1; the associations remained statistically highly significant ($P < 5.0 \times 10^{-8}$; **Supplementary Table 4**).

rs11571833 and rs56084662, localizing to 13q13.1 near or within *BRCA2*, are rare (MAF < 0.01), map 103 kb apart (32,972,376 bp and 32,869,614 bp, respectively) and are moderately correlated (r^2 = 0.45 and D' = 0.82 based on genotypes from the Heidelberg-EPIC, IARC, ICR and Toronto replication series; **Fig. 3**). rs11571833 (c.9976A>T) is responsible for BRCA2 p.Lys3326X, whereas rs56084662 is located in the 3′ UTR of *FRY*. Although the association provided by rs11571833 was substantially stronger than that provided by rs56084662 in the combined analysis (OR = 1.83, $P = 2.11 \times 10^{-19}$ and $P = 1.88 \times 10^{-15}$, respectively), a conditional analysis based on directly genotyped samples in the replication series was consistent with the two SNPs tagging the same haplotype. The association at rs11571833 is driven primarily by a relationship with SQ histology rather than AD histology (OR = 2.47, $P = 4.74 \times 10^{-20}$ and OR = 1.47, $P = 4.66 \times 10^{-4}$, respectively; **Fig. 2** and **Supplementary Table 3**). A stronger role for *BRCA2* in SQ etiology than in AD etiology is reflected in the higher observed mutational frequency in the respective lung cancers (~6% and 1% (refs. 18,19)). Thr9976 was recently shown to confer a 1.26-fold increased breast cancer risk[20] and has been suggested previously as a risk factor for esophageal and pancreatic cancers[21,22]. We found no evidence for an association between Thr9976 and lung cancer risk in nonsmokers using directly genotyped samples (**Supplementary Table 2**); however, these cases comprised <10% of each cohort, and therefore our power to demonstrate a relationship was limited. Previous analyses of families carrying highly penetrant *BRCA2* mutations have found either no evidence for any excess risk or a reduced risk of lung cancer in carriers[23,24]. A possible explanation for these observations is that members of the families studied tended to smoke less than the general population[24].

The RAD51-BRCA2 interaction is pivotal for BRCA2-mediated double strand–break repair, and exon 27 of *BRCA2* encodes one of the highly conserved RAD51 binding domains: homozygous deletion of exon 27 in mice confers susceptibility to tumors, including lung cancer[25]. Thr9976 leads to the loss of the C-terminal domain of BRCA2, inviting speculation that the SNP is functional. Although the deleted region is distal to the RAD51 binding domain and an impact on nuclear localization is unknown[26,27], the nearby BRCA2 p.Thr3387Ala alteration interrupts CHK2 phosphorylation and abrogates BRCA2-CHK2-RAD51–mediated recombination repair[28]. Alternatively, the association might be a consequence of linkage disequilibrium (LD) with another *BRCA2* mutation. Studies of families with breast cancer of northern European ancestry show that the *BRCA2* c.6275delTT and c.4889C>G mutations, which are highly

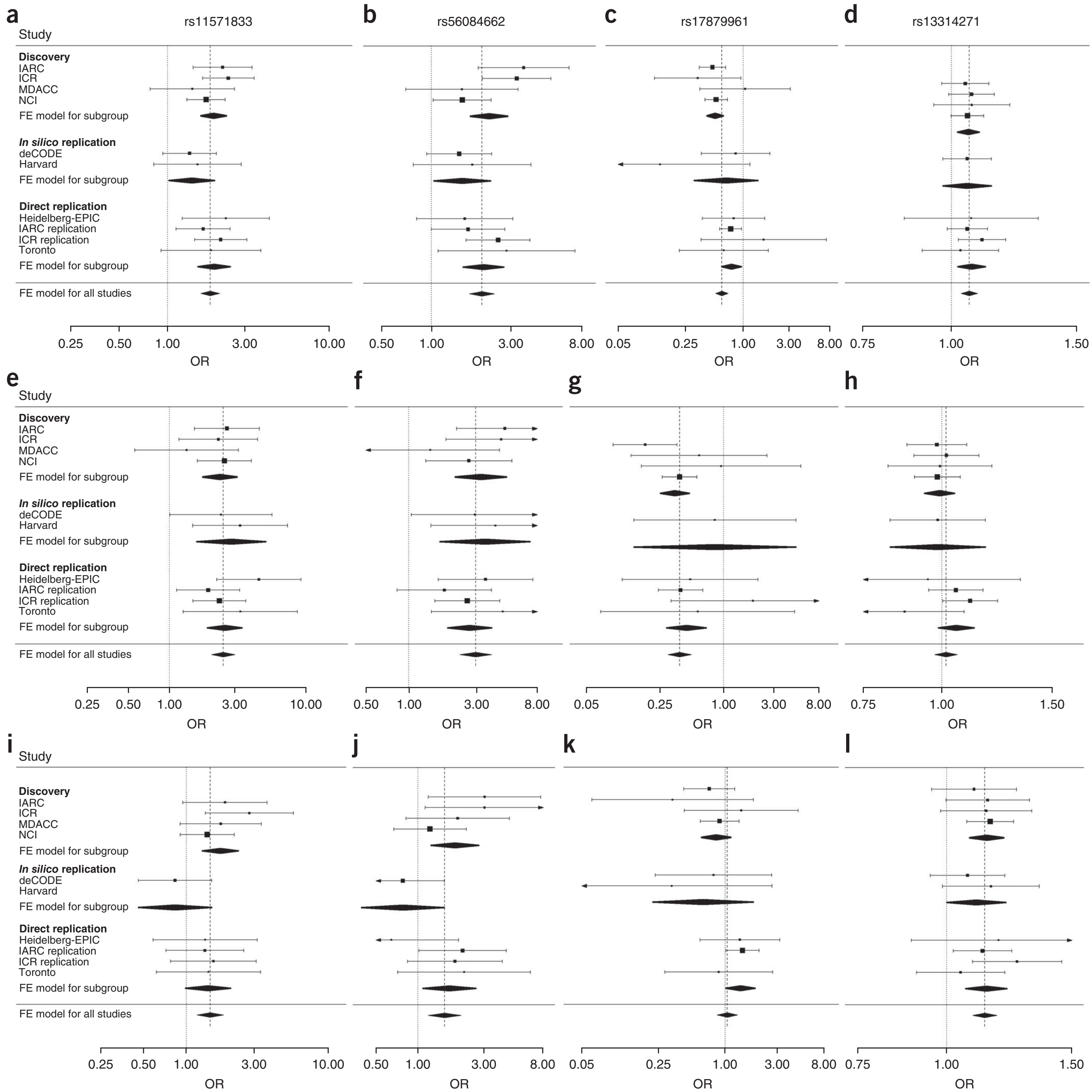


Figure 2 Plots of the ORs of lung cancer associated with 13q13.1 (rs11571833 and rs56084662), 22q12.1 (rs17879961) and 3q28 (rs13314271) risk loci. (**a**–**l**) All lung cancer based on 21,594 lung cancer cases and 54,156 controls (**a**–**d**), SQ based on 6,477 SQ cases and 53,333 controls (**e**–**h**) and AD based on 7,031 AD cases and 53,189 controls (**i**–**l**). The studies are weighted according to the inverse of the variance of the log of the OR calculated by unconditional logistic regression. Horizontal lines indicate the 95% CIs. Boxes are the OR point estimates, and the area of the box is proportional to the weight of the study. Diamonds and broken lines indicate the overall summary estimate derived under a fixed-effects (FE) model, with the CI given by the width. Unbroken vertical lines show the null value (OR = 1.0).

penetrant for breast and ovarian cancer, originated on a p.Lys3326X haplotype[29]. To gain further insight into a probable genetic basis of the 13q13.1 lung cancer association, we sequenced germline DNA from 70 individuals with lung cancer who carried c.9976A>T from the UK Genetic Lung Cancer Predisposition Study for the c.6275delTT and c.4889C>G mutations; we did not find c.6275delTT or c.4889C>G in any of these individuals. Similarly, sequencing the coding region of *BRCA2* identified no clearly pathogenic mutations among 13 individuals from the 1958 British Birth Cohort (58BC), 11 individuals with lung cancer from IARC or 24 individuals with lung cancer carrying Thr9976 from TCGA. In Iceland, Thr9976 is not correlated with the founder *BRCA2* mutation resulting in p.256_257del (c.999del5), which greatly increases the risk of breast and ovarian cancer. Paradoxically, whereas Thr9976 is a risk factor for lung cancer, in this population this SNP is not associated with risk of breast or ovarian cancer (**Supplementary Table 5**). Although *in vitro* studies

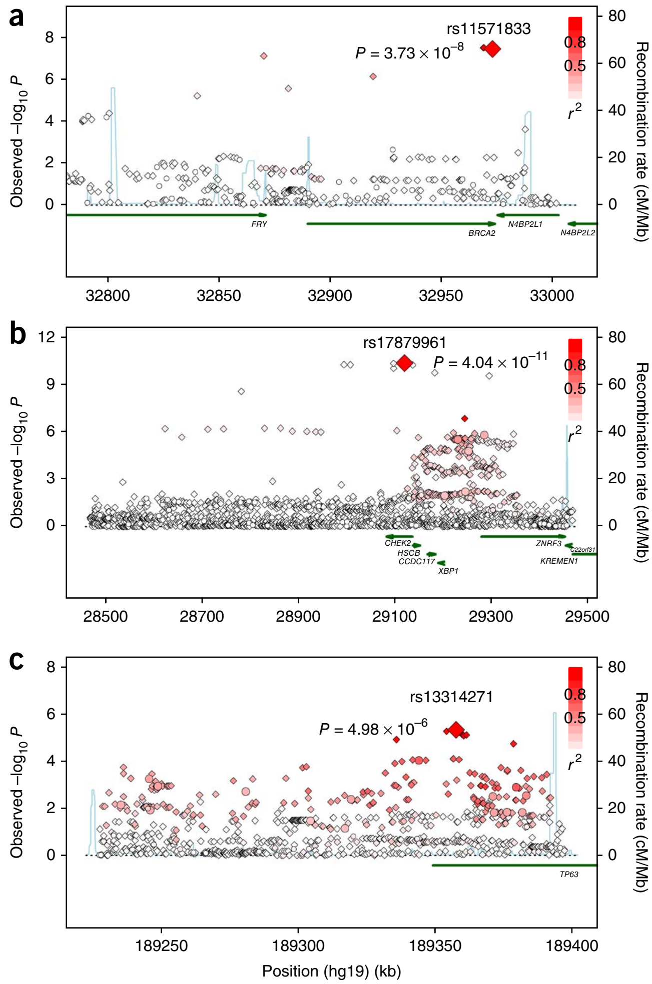


Figure 3 Regional plots of associations at susceptibility loci for SQ and AD. (**a**–**c**) Association results and recombination rates for the 13q13.1 in SQ (**a**), 22q12.1 in SQ (**b**) and 3q28 in AD (**c**). The SQ-related plots (**a**,**b**) were based on 3,275 SQ cases and 15,038 controls from the discovery phase; the AD-related plot (**c**) was based on 3,442 AD cases and 14,894 controls from the discovery phase. Association results of both genotyped (circles) and imputed (diamonds) SNPs in the GWAS samples and recombination rates for each locus are shown. For each plot, $-\log_{10} P$ values (*y* axes) of the SNPs are shown according to their chromosomal positions (*x* axes). The top genotyped SNP in each combined analysis is indicated by a large diamond and is labeled by its rsID. The color intensity of each symbol reflects the extent of LD with the top genotyped SNP: white ($r^2 = 0$) through to dark red ($r^2 = 1.0$). Genetic recombination rates (cM/Mb), estimated using HapMap CEU samples, are shown with a light blue line. Physical positions are based on NCBI build 37 of the human genome. Also shown are the relative positions of genes and transcripts mapping to each region of association. Genes have been redrawn to show the relative positions; therefore, maps are not to physical scale.

have failed to demonstrate that p.Lys3326X affects DNA repair[30], our findings raise the possibility that p.Lys3326X may have a direct effect on lung cancer risk. The fact that somatic mutation of *BRCA2* is not associated with p.Lys3326X carrier status[19] (**Supplementary Table 6a**) suggests that any impact the SNP has on lung cancer risk is mediated through alternative mechanisms.

The relationship at 22q12.1 between the rs17879961 (c.470T>C) and SQ in the combined series (OR = 0.38, $P = 1.27 \times 10^{-13}$) validates an association that has been reported previously[31,32] (**Fig. 2** and **Supplementary Tables 3** and **4**). The frequency of rs17879961 varies markedly between populations: it has a MAF of ~5% in eastern Europeans (for example, individuals in the IARC series) but is almost monomorphic in most northern Europeans. This likely accounts for the failure to demonstrate a significant relationship in the ICR, MDACC, Toronto and deCODE series, which comprise largely western European populations (**Fig. 2** and **Supplementary Table 3**). rs17879961 is responsible for the missense mutation in *CHEK2* resulting in p.Ile157Thr; *CHEK2* is a cell cycle–control gene encoding a pluripotent kinase that can cause arrest or apoptosis in response to DNA damage. Acquired mutation of *CHEK2* is rarely seen in lung cancer, and the CHEK2 p.Ile157Thr alteration does not appear to correlate with mutation (**Supplementary Table 6a**), raising the possibility that carrier status *per se* influences cancer risk. The p.Ile157Thr substitution lies in a functionally important domain of CHEK2 and causes reduced or abolished binding of principal substrates. Although Cys470 increases breast cancer risk[33], here Cys470 was associated with reduced lung cancer risk. A mechanism for the paradoxical associations is not immediately apparent. However, *CHEK2* can have opposite effects on damaged stem cells, retarding stem cell division until DNA damage is repaired or activating apoptosis if damage cannot be repaired. Although speculative, in the presence of continued DNA damage to squamous epithelia by tobacco smoke, the normal stem cell defenses involving CHEK2 might be attenuated by a reduction in CHEK2 activity as a result of p.Ile151Thr[31]. Concordant with such a model is our observation of a paradoxically increased lung cancer risk in nonsmokers ($P = 0.05$) and in correlated subgroups of AD and women, although this increase was based on small numbers (**Supplementary Table 2**).

The association between variation at 3q28 marked by rs13314271 and lung cancer risk was restricted to AD (OR = 1.13, $P = 7.22 \times 10^{-10}$; **Fig. 2** and **Supplementary Table 3**). rs13314271 maps within intron 1 of *TP63* (**Fig. 3**). Variation at *TP63* defined by the intron 1 SNP rs4488809, which is in complete LD with rs13314271 ($r^2 = 1.00$, $D' = 1.00$), is associated with AD in Asians[8]. Our findings provide robust evidence for the generalizability of a relationship between 3q28 variation and AD. We found a weak association between rs13314271 and lung cancer risk in nonsmokers ($P = 0.03$; **Supplementary Table 2b**). *TP63* is a member of the tumor suppressor *TP53* gene family, which is pivotal in cellular differentiation and responsiveness to cellular stress[34,35]. Exposure of cells to DNA damage leads to induction of *TP63*, and both isoforms have the ability to transactivate *TP53* target genes, thereby affecting cellular responsiveness to DNA damage[36]. Although rs13314271 does not map to an evolutionary conserved region, rs7636839, which is correlated with rs13314271 and rs4488809 ($r^2 = 1.0$), does map to an evolutionarily conserved region and has predicted enhancer activity (**Supplementary Table 6b**). Moreover, rs4488809 has been shown to be an expression quantitative trait locus for *TP63* in lung tissue[37]. Although the mechanism by which 3q28 variation affects AD development is unknown, accumulation of DNA damage and a lack of response to genotoxic stress are recognized to contribute to lung carcinogenesis; hence, loss of repair fidelity as a consequence of differential *TP63* expression is likely deleterious.

There was no association between rs11571833, rs17879961 and rs13314271 genotypes and cigarette consumption on the basis of smoking information on 43,693 Icelandic subjects (**Supplementary Table 7**), which is in contrast to the association of 15q25 and risk of lung cancer.

Although there is some overlap, distinct DNA lesions are ostensibly repaired by different DNA repair pathways. Histology-specific relationships seen implicate the BRCA2-CHEK2-RAD52 double strand–break repair and homologous recombination pathways as a determinant of SQ and defective TP53 and TERT apoptosis-telomerase regulation as a basis of AD risk.

In conclusion, our findings provide further evidence for inherited genetic susceptibility to lung cancer and underscore the importance of searching for histology-specific risk variants. Our data also provide an important proof of principle that 1000 Genomes imputation can be used to detect new, low-frequency, large-effect associations, thereby extending the utility of preexisting GWAS data. Notably, this study facilitated the identification of BRCA2 Thr9976, which is the strongest genetic association in lung cancer reported so far. For a smoker carrying this variant (2% of the population), the risk of developing lung cancer is approximately doubled, which may have implications for identifying high-risk ever-smoking subjects for lung cancer screening. Additionally, future study of the effects of PARP inhibition in smokers with lung cancer carrying BRCA2 Thr9976 may be warranted.

URLs. R suite, http://www.r-project.org/; 1000 Genomes Project, http://www.1000genomes.org/; SNAP, http://www.broadinstitute.org/mpg/snap/; IMPUTE2, http://mathgen.stats.ox.ac.uk/impute/impute_v2.html; MACH, http://www.sph.umich.edu/csg/abecasis/MACH/; Minimac, http://genome.sph.umich.edu/wiki/Minimac/; SNPTEST, https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html; ProbABEL, http://www.genabel.org/packages/ProbABEL; mach2dat, http://genome.sph.umich.edu/wiki/Mach2dat:_Association_with_MACH_output; Wellcome Trust Case Control Consortium, http://www.wtccc.org.uk/; RegulomeDB, http://regulome.stanford.edu/; HaploReg v2, http://www.broadinstitute.org/mammals/haploreg/haploreg.php; Transdisciplinary Research In Cancer of the Lung (TRICL), http://u19tricl.org/; Genetic Associations and MEchanisms in ONcology (GAME-ON) consortium, http://epi.grants.cancer.gov/gameon/; International Lung Cancer Consortium (ILCCO), http://ilcco.iarc.fr/; Icelandic Cancer Registry, http://www.krabbameinsskra.is/; Genome Analysis Toolkit (GATK), http://www.broadinstitute.org/gatk/; The Cancer Genome Atlas (TCGA), http://cancergenome.nih.gov/; Leiden Open Variation Database (LOVD), http://www.lovd.nl/3.0/home/; Breast Cancer IARC database, http://brca.iarc.fr/.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank all individuals who participated in this study. We are also grateful to the patients, clinicians and allied health care professions. We thank Z. Chen and K. Boyle for sample handling and data management of the Toronto study, and L. Admas and L.R. Zhang for field recruitment. We thank L. Su, Y. Zhao, G. Liu, J. Wain, R. Heist and K. Asomaning for providing computing support at MDACC. We thank G. Thomas and Synergy Lyon Cancer (Lyon France) for high performance computing support and J. Olivier and A. Chabrier for IARC's PGM ion torrent sequencing optimization and TaqMan genotyping, respectively. We thank D. Goldgar for sharing information from The Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA) on sequence variation in BRCA2 from familial breast cancer analysis. We acknowledge the Icelandic Cancer Registry (http://www.krabbameinsskra.is/indexen.jsp?id=summary) for assistance in the ascertainment of the Icelandic patients with lung cancer. The ICR study made use of genotyping data from the Wellcome Trust Case-Control Consortium 2 (WTCCC2); a full list of the investigators who contributed to the generation of the data is available from http://www.wtccc.org.uk. We acknowledge The Cancer Genome Atlas (TCGA) for their contribution of lung cancer genomic data to this study (TCGA Project Number 3230). We also acknowledge support from the National Institute for Health Research Biomedical Research Centre at the Royal Marsden Hospital. This study was supported by the NIH (U19CA148127, R01CA055769, 5R01CA127219, 5R01CA133996 and 5R01CA121197). The work performed at ICR was supported by Cancer Research UK (C1298/A8780 and C1298/A8362), National Cancer Research Network (NCRN), HEAL, Sanofi-Aventis and National Health Service funding to the Royal Marsden Hospital and Institute of Cancer Research, as well as the National Institute for Health Research Biomedical Research Centre. B.K. was the recipient of a Sir John Fisher Foundation PhD studentship. Work at ICR was also supported by NIH GM103534 and the Institute for Quantitative Biomedical Sciences at Dartmouth to C.I.A. The work performed in Toronto was supported by The Canadian Cancer Society Research Institute (020214), Ontario Institute of Cancer and Cancer Care Ontario Chair Award to R.J.H. and G.L. and the Alan Brown Chair and Lusi Wong Programs at the Princess Margaret Hospital Foundation. The work performed at Heidelberg was supported by Deutsche Krebshilfe (70-2387 and 70-2919) and the German Federal Ministry of Education and Research (EPIC-Heidelberg). The work performed at IARC was supported by the Institut National du Cancer, France, the European Community (LSHG-CT-2005-512113), the Norwegian Cancer Association, the Functional Genomics Programme of Research Council of Norway, the European Regional Development Fund and the State Budget of the Czech Republic (RECAMO, CZ.1.05/2.1.00/03.0101), the NIH (R01-CA111703 and UO1-CA63673), the Fred Hutchinson Cancer Research Center, the US NCI (R01 CA092039), an FP7 grant (REGPOT 245536), the Estonian Government (SF0180142s08), the EU European Regional Development Fund in the frame of Centre of Excellence in Genomics and Estonian Research Infrastructure's Roadmap and the University of Tartu (SP1GVARENG) and an IARC Postdoctoral Fellowship (M.N.T.). Work at the NCI was supported by the Intramural Research Program of the NIH, the NCI, US Public Health Service contracts NCI (N01-CN-45165, N01-RC-45035, N01-RC-37004, NO1-CN-25514, NO1-CN-25515, NO1-CN-25512, NO1-CN-25513, NO1-CN-25516, NO1-CN-25511, NO1-CN-25524, NO1-CN-25518, NO1-CN-75022, NO1-CN-25476 and NO1-CN-25404), the American Cancer Society, the NIH Genes, Environment and Health Initiative in part by HG-06-033-NCI-01 and RO1HL091172-01, genotyping at the Johns Hopkins University Center for Inherited Disease Research (U01HG004438 and NIH HHSN268200782096C) and study coordination at the GENEVA Coordination Center (U01 HG004446). Work was also supported by NIH grants (P50 CA70907, R01CA121197, RO1 CA127219, U19 CA148127 and RO1 CA55769) and a Cancer Prevention Research Institute of Texas grant (RP100443). Genotyping was provided by the Center for Inherited Disease Research (CIDR). Work performed at Harvard was supported by the NIH (CA074386, CA092824 and CA090578). The Icelandic study was supported in part by NIH DA17932.

AUTHOR CONTRIBUTIONS

R.S.H. and Y. Wang conceived the study and provided overall project management and drafted the paper. In the UK, Y. Wang performed statistics and bioinformatics of UK data and conducted all meta-analyses; additional support was provided by M.H.; P. Broderick oversaw genotyping and sequencing; A.L. and B.K. performed genotyping and Sanger sequencing; A. Matakidou, T.E. and R.S.H. were responsible for the development and operation of the Genetic Lung Cancer Predisposition Study (GELCAPS); and D.C. and P. Broderick performed next-generation sequencing. At IARC, J.D.M. and P. Brennan provided overall project management; M.N.T., M.D.-S., V.G. and M.V. performed statistics and bioinformatics of IARC data and conducted meta-analysis; J.D.M. and F.L.-K. oversaw genotyping and sequencing; and G.S., D.Z., N.S.-D., J. Lissowska, P.R., E.F., D.M., V.B., L.F., V.J., H.E.K., M.E.G., F.S., L.V., I.N., C.C., G.G., M. Lathrop, S.B., T.V., K.V., M.N., A. Metspalu, M. Lathrop, J. Lubiński, Mattias Johansson, P.V., A.A., F.C.-C., H.B.-d.-M., D.T., K.-T.K., Mikael Johansson, E.W., A.T., R.K. and E.R. provided samples and data. For the Dartmouth and MDACC component, C.I.A. provided overall project management, obtained support for genotyping and contributed to statistical analyses; W.V.C. performed imputation analysis; Y.H. performed statistical analyses; and M.R.S. oversaw sample collection and development of the epidemiological studies. M.R.S. was also responsible for collecting samples that are a part of this research. X.W. provided ongoing support for the research protocol and supported large laboratory management of samples. Y.Y. and J.G. performed genotyping. At the NCI, M.T.L. was responsible for the overall project and managed the Environment and Genetics in Lung Cancer Etiology (EAGLE) study; N.E.C. managed the Prostate, Lung, Colon, Ovary Screening Trial (PLCO) study; D.A. managed the α-Tocopherol, β-Carotene Cancer Prevention Study (ATBC); S.M.G. and V.L.S. managed the Cancer Prevention Study II Nutrition Cohort (CPS-II) study; N.C. and W.W. performed statistical analyses; Z.W. performed genotyping and imputation analysis; and S.J.C. oversaw genotyping and imputation analysis. At decode, T.R. and K.S. were responsible for the development and operation of deCODE's lung cancer study; and G.T. and P.S. performed the imputations and statistical analysis of the Icelandic data. At Harvard, D.C.C. was responsible for the overall conduct of the project; L.S. was responsible for sample management, genotyping and laboratory quality control; and Y. Wei performed data management and statistical analyses. For the Heidelberg-EPIC replication, M. Laplana managed DNA samples and performed genotyping; A. Rosenberger managed genotype and phenotype information; A. Risch supervised genotyping and data analysis; and R.K., A. Risch and H.D. conceived and managed studies that contributed samples. For the Toronto replication, R.J.H. and G.L. provided

overall supervision of the study conduct, including study design, field recruitment, genotyping and statistical analysis; and X.Z. performed the statistical analysis.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Ferlay, J. *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* **127**, 2893–2917 (2010).
2. Hung, R.J. *et al.* A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* **452**, 633–637 (2008).
3. Amos, C.I. *et al.* Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat. Genet.* **40**, 616–622 (2008).
4. Thorgeirsson, T.E. *et al.* A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* **452**, 638–642 (2008).
5. McKay, J.D. *et al.* Lung cancer susceptibility locus at 5p15.33. *Nat. Genet.* **40**, 1404–1406 (2008).
6. Wang, Y. *et al.* Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat. Genet.* **40**, 1407–1409 (2008).
7. Hu, Z. *et al.* A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. *Nat. Genet.* **43**, 792–796 (2011).
8. Miki, D. *et al.* Variation in *TP63* is associated with lung adenocarcinoma susceptibility in Japanese and Korean populations. *Nat. Genet.* **42**, 893–896 (2010).
9. Lan, Q. *et al.* Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. *Nat. Genet.* **44**, 1330–1335 (2012).
10. Travis, W.D. *et al.* International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society: international multidisciplinary classification of lung adenocarcinoma: executive summary. *Proc. Am. Thorac. Soc.* **8**, 381–385 (2011).
11. Broderick, P. *et al.* Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. *Cancer Res.* **69**, 6633–6641 (2009).
12. Landi, M.T. *et al.* A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. *Am. J. Hum. Genet.* **85**, 679–691 (2009).
13. Timofeeva, M.N. *et al.* Influence of common genetic variation on lung cancer risk: meta-analysis of 14,900 cases and 29,485 controls. *Hum. Mol. Genet.* **21**, 4980–4995 (2012).
14. Shi, J. *et al.* Inherited variation at chromosome 12p13.33, including *RAD52*, influences the risk of squamous cell lung carcinoma. *Cancer Discov.* **2**, 131–139 (2012).
15. Huang, Y.T. *et al.* Cigarette smoking increases copy number alterations in nonsmall-cell lung cancer. *Proc. Natl. Acad. Sci. USA* **108**, 16345–16350 (2011).
16. Rafnar, T. *et al.* Sequence variants at the *TERT-CLPTM1L* locus associate with many cancer types. *Nat. Genet.* **41**, 221–227 (2009).
17. Mathieson, I. & McVean, G. Differential confounding of rare and common variants in spatially structured populations. *Nat. Genet.* **44**, 243–246 (2012).
18. Imielinski, M. *et al.* Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107–1120 (2012).
19. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525 (2012).
20. Michailidou, K. *et al.* Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat. Genet.* **45**, 353–361 (2013).
21. Akbari, M.R. *et al.* Germline *BRCA2* mutations and the risk of esophageal squamous cell carcinoma. *Oncogene* **27**, 1290–1296 (2008).
22. Martin, S.T. *et al.* Increased prevalence of the *BRCA2* polymorphic stop codon K3326X among individuals with familial pancreatic cancer. *Oncogene* **24**, 3652–3656 (2005).
23. Breast Cancer Linkage Consortium. Cancer risks in *BRCA2* mutation carriers. *J. Natl. Cancer Inst.* **91**, 1310–1316 (1999).
24. van Asperen, C.J. *et al.* Cancer risks in *BRCA2* families: estimates for sites other than breast and ovary. *J. Med. Genet.* **42**, 711–719 (2005).
25. McAllister, K.A. *et al.* Cancer susceptibility of mice with a homozygous deletion in the COOH-terminal domain of the *Brca2* gene. *Cancer Res.* **62**, 990–994 (2002).
26. Spain, B.H., Larson, C.J., Shihabuddin, L.S., Gage, F.H. & Verma, I.M. Truncated BRCA2 is cytoplasmic: implications for cancer-linked mutations. *Proc. Natl. Acad. Sci. USA* **96**, 13920–13925 (1999).
27. Yano, K. *et al.* Nuclear localization signals of the BRCA2 protein. *Biochem. Biophys. Res. Commun.* **270**, 171–175 (2000).
28. Bahassi, E.M. *et al.* The checkpoint kinases Chk1 and Chk2 regulate the functional associations between hBRCA2 and Rad51 in response to DNA damage. *Oncogene* **27**, 3977–3985 (2008).
29. Mazoyer, S. *et al.* A polymorphic stop codon in *BRCA2*. *Nat. Genet.* **14**, 253–254 (1996).
30. Wu, K. *et al.* Functional evaluation and cancer risk assessment of *BRCA2* unclassified variants. *Cancer Res.* **65**, 417–426 (2005).
31. Brennan, P. *et al.* Uncommon *CHEK2* mis-sense variant and reduced risk of tobacco-related cancers: case control study. *Hum. Mol. Genet.* **16**, 1794–1801 (2007).
32. Cybulski, C. *et al.* Constitutional *CHEK2* mutations are associated with a decreased risk of lung and laryngeal cancers. *Carcinogenesis* **29**, 762–765 (2008).
33. Han, F.F., Guo, C.L. & Liu, L.H. The effect of *CHEK2* variant I157T on cancer susceptibility: evidence from a meta-analysis. *DNA Cell Biol.* **32**, 329–335 (2013).
34. Flores, E.R. The roles of p63 in cancer. *Cell Cycle* **6**, 300–304 (2007).
35. Katoh, I., Aisaki, K.I., Kurata, S.I., Ikawa, S. & Ikawa, Y. p51A (TAp63γ), a p53 homolog, accumulates in response to DNA damage for cell regulation. *Oncogene* **19**, 3126–3130 (2000).
36. Petitjean, A. *et al.* Properties of the six isoforms of p63: p53-like regulation in response to genotoxic stress and cross talk with ΔNp73. *Carcinogenesis* **29**, 273–281 (2008).
37. Hao, K. *et al.* Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS Genet.* **8**, e1003029 (2012).

[1]Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, Surrey, UK. [2]International Agency for Research on Cancer (IARC, World Health Organization (WHO)), Lyon, France. [3]deCODE Genetics, Amgen, Reykjavik, Iceland. [4]Department of Health and Human Services, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health (NIH), Bethesda, Maryland, USA. [5]Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Ontario, Canada. [6]Division of Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Heidelberg, Germany. [7]Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts, USA. [8]Center for Genomic Medicine, Department of Community and Family Medicine, Geisel School of Medicine, Dartmouth College, Lebanon, New Hampshire, USA. [9]Information Management Services, Inc., Rockville, Maryland, USA. [10]Princess Margaret Hospital, University Health Network, Toronto, Ontario, Canada. [11]Division of Cancer Epidemiology, DKFZ, Heidelberg, Germany. [12]Translational Lung Research Center Heidelberg (TLRC-H), Member of the German Center for Lung Research (DZL), Heidelberg, Germany. [13]Epidemiology Research Program, American Cancer Society, Atlanta, Georgia, USA. [14]Department of Genetics, University of Texas MD Anderson Cancer Center, Houston, Texas, USA. [15]Institute of Carcinogenesis, Russian N.N. Blokhin Cancer Research Centre, Moscow, Russia. [16]Department of Epidemiology, Institute of Occupational Medicine, Lodz, Poland. [17]The M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland. [18]National Institute of Environmental Health, Budapest, Hungary. [19]Regional Authority of Public Health, Banská Bystrica, Slovak Republic. [20]National Institute of Public Health, Bucharest, Romania. [21]1st Faculty of Medicine, Institute of Hygiene and Epidemiology, Charles University in Prague, Prague, Czech Republic. [22]Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic. [23]Palacky University, Olomouc, Czech Republic. [24]Department of Laboratory Medicine, Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, Norway. [25]Department of Laboratory Medicine, Children's and Women's Health, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway. [26]Department of Public Health and General Practice, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway. [27]Department of Community Medicine, University of Tromsø, Tromsø, Norway. [28]Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [29]INSERM U946, Paris, France. [30]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. [31]Department of Biomedicine, University of Bergen, Bergen, Norway. [32]Estonian Genome Center, Institute of Molecular and Cell Biology, Tartu, Estonia. [33]Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland. [34]Department of Genetics and Pathology, International Hereditary Cancer Center, Pomeranian Medical University, Szczecin, Poland. [35]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College, London, UK. [36]Unit of Molecular and Genetic Epidemiology, HuGeF Foundation, Torino, Italy. [37]Unit of Nutrition, Environment and Cancer, Cancer Epidemiology Research Program, Catalan Institute of Oncology, Barcelona, Spain. [38]INSERM, Centre for Research in Epidemiology and Population Health (CESP), U1018, Nutrition, Hormones and Women's Health Team, Villejuif, France. [39]Université Paris Sud, UMRS 1018, Villejuif, France. [40]Institut Gustave Roussy, Villejuif, France. [41]National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands. [42]Department of Gastroenterology and Hepatology, University Medical Centre, Utrecht, The Netherlands. [43]Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA. [44]Bureau of Epidemiologic Research, Academy of Athens, Athens, Greece. [45]Hellenic Health Foundation, Athens, Greece. [46]University of Cambridge School of Clinical Medicine, Clinical Gerontology Unit, Addenbrooke's Hospital, Cambridge, UK. [47]Department of Radiation Sciences, Umeå Universitet, Umeå, Sverige, Sweden. [48]Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway. [49]Department of Research, Cancer Registry of Norway, Oslo, Norway. [50]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [51]Samfundet Folkhälsan, Helsinki, Finland. [52]Danish Cancer Society Research Center, Copenhagen, Denmark. [53]Centre d'Etude du Polymorphisme Humain (CEPH), Paris, France. [54]Department of Epidemiology, University of Texas MD Anderson Cancer Center, Houston, Texas, USA. [55]Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, Texas, USA. [56]Department of Thoracic Surgery, Thoraxklinik at University Hospital Heidelberg, Heidelberg, Germany. [57]Department of Genetic Epidemiology, University of Göttingen, Göttingen, Germany. [58]Cancer Research UK Cambridge Institute, Li Ka Shing Centre, Cambridge, UK. [59]Department of Oncology, Cambridge University Hospitals National Health Service Foundation Trust, Cambridge, UK. [60]Addenbrooke's Hospital, Cambridge Biomedical Campus, Cambridge, UK. [61]These authors contributed equally to this work. [62]These authors jointly directed this work. Correspondence should be addressed to R.S.H. (richard.houlston@icr.ac.uk) or M.T.L. (landim@mail.nih.gov).

ONLINE METHODS

Studies. The study was conducted under the auspices of the Transdisciplinary Research In Cancer of the Lung (TRICL) Research Team, which is a part of the Genetic Associations and MEchanisms in ONcology (GAME-ON) consortium and is associated with the International Lung Cancer Consortium (ILCCO). Tumors from patients were classified as AD, SQ, large-cell carcinoma (LCC), mixed adenosquamous carcinoma (MADSQ) and other NSCLC histologies following either the International Classification of Diseases for Oncology (ICD-O) or WHO coding. Tumors with overlapping histologies were classified as mixed.

Ethics. All participants provided informed written consent. All studies were reviewed and approved by institutional ethics review committees at the involved institutions.

GWAS. The meta-analysis was based on data from four previously reported lung cancer GWAS of European populations: the MDACC GWAS[3], the ICR GWAS[6], the NCI GWAS[13] and the IARC GWAS[2]. In each of the studies, SNP genotyping had been performed using Illumina HumanHap 317, 317+240S, 370, 550, 610 or 1M arrays (**Supplementary Table 1**).

IARC GWAS. The IARC GWAS[2] comprised 3,062 lung cancer cases and 4,455 controls derived from five case-control studies: (i) the Carotene and Retinol Efficacy Trial (CARET) cohort[38]; (ii) the Central Europe multicenter hospital-based case-control study[39,40]; (iii) the hospital-based case-control study from France[40]; (iv) the hospital-based case-control lung cancer study from Estonia[41,42]; and (v) the population-based HUNT2/Tromsø IV lung cancer studies[43]. Patient and control DNAs were derived from EDTA–venous blood samples. The patients with lung cancer were classified according to ICD-O-3: SQ: 8070/3, 8071/3, 8072/3, 8074/3; AD: 8140/3, 8250/3, 8260/3, 8310/3, 8480/3, 8560/3, 8251/3, 8490/3, 8570/3, 8574/3; with tumors with overlapping histologies being classified as mixed. After applying standardized quality-control procedures, 2,533 cases and 3,791 controls were included in the current analysis (**Supplementary Table 1**).

NCI GWAS. Details of the NCI GWAS have been reported previously. Briefly, the study comprised samples from four series: (i) the Environment and Genetics in Lung cancer Etiology (EAGLE) study, a population-based case-control study of 2,100 lung cancer cases and 2,120 healthy controls enrolled in Italy between 2002 and 2005 (ref. 44), in which cancers were classified according to the ICD-O coding for histology and grading and histology of ~10% of tumors was confirmed by an independent pathologist from the NCI; (ii) the α-Tocopherol, β-Carotene Cancer Prevention Study (ATBC), a randomized primary prevention trial of 29,133 male smokers enrolled in Finland between 1985 and 1993 (ref. 45), in which ICD-O-2 and ICD-O-3 were used to classify tumors and cases diagnosed between 1985 and 1999 had histology reviewed by at least one pathologist (after 1999, histological coding (ICD-O-2 and ICD-O-3) was derived from the Finnish Cancer Registry); (iii) the Prostate, Lung, Colon, Ovary Screening Trial (PLCO), a randomized trial of 150,000 individuals enrolled in 10 US study centers between 1992 and 2001 (ref. 46), in which ICD-O-2 was used to classify tumors and quality assurance measures included reabstraction of 50 lung cancer diagnoses per year; and (iv) the Cancer Prevention Study II Nutrition Cohort (CPS-II), a cohort study of approximately 184,000 individuals enrolled by the American Cancer Society between 1992 and 1993 in 21 US states, of which 109,379 provided a blood (36%) or buccal (64%) sample between 1998 and 2003 (refs. 12,47) and tumor histology was abstracted from Certified Tumor Registrars and coded using WHO ICD-O-2 and ICD-O-3. In this study, quality assurance was done by reabstracting 10% of all cancer diagnoses per year. After initial data quality control, the NCI GWAS included 5,739 cases and 5,848 controls; however, an additional 26 cases and 112 controls were excluded because of changes in case status and further quality-control filtering. The current meta-analysis included 5,713 lung cancer cases and 5,736 controls from the NCI GWAS (**Supplementary Table 1**).

ICR GWAS. The ICR GWAS comprised 1,952 cases (1,166 male; mean age at diagnosis 57 years, s.d. 6 years) with pathologically confirmed lung cancer ascertained through the Genetic Lung Cancer Predisposition Study (GELCAPS) conducted between March 1999 and July 2004 (ref. 48). All cases were British residents and were self-reported to be of European ancestry. To ensure that data and samples were collected from *bona fide* lung cancer cases and avoid issues of bias from survivorship, only incident cases with histologically or cytologically (if not AD) confirmed primary disease were ascertained. Tumors from patients were classified according to ICD-O3: specifically, SQ: 8070/3, 8071/3, 8072/3, 8074/3; AD: 8140/3, 8250/3, 8260/3, 8310/3, 8480/3, 8560/3, 8251/3, 8490/3, 8570/3, 8574/3; with tumors with overlapping histologies being classified as mixed. Patient DNA was derived from EDTA–venous blood samples using conventional methodologies. Genotype frequencies were compared with publicly accessible data generated by the UK Wellcome Trust Case-Control Consortium 2 (WTCCC2) study[49] of individuals from the 1958 British Birth Cohort (58BC), and blood service was typed using Illumina Human1.2M-Duo Custom_v1 Array BeadChips.

MDACC GWAS. Cases and controls were ascertained from a case-control study at the University of Texas MD Anderson Cancer Center conducted between 1997 and 2007 (ref. 3). Cases were newly diagnosed patients with histologically confirmed lung cancer presenting at MD Anderson Cancer who had not previously received treatment other than surgery. Clinical and pathological data were abstracted from patient medical records, and lung cancer histology was coded according to the major histological groups. Specifically, as per ICD-O-2, these groups were SQ: 8070/3; AD: 8140/3, 8250/3, 8260/3, 8310/3, 8480/3, 8251/3, 8490/3. Only patients with predominantly or wholly AD or SQ cancers were included; those with mixed histology or unspecified lung cancers were excluded from the study. Controls were healthy individuals seen for routine care at Kelsey-Seybold clinics in the Houston metropolitan area. Controls were frequency matched to cases according to smoking behavior, age in 5-year categories, ethnicity and sex. Former smoking controls were further frequency matched to former smoking cases according to the number of years since smoking cessation (in 5-year categories). After applying quality controls, data were available on 1,150 cases and 1,134 controls.

Quality control of GWAS data sets. Standard quality control was performed on all scans, excluding individuals with low call rate (<90%) and extremely high or low heterozygosity ($P < 1.0 \times 10^{-4}$), as well as all individuals evaluated to be of non-European ancestry (using the HapMap version 2 CEU, JPT/CHB and YRI populations as a reference; **Supplementary Table 1**). For apparent first-degree relative pairs, we removed the control from a case-control pair; otherwise, we excluded the individual with the lower call rate.

Replication series. To validate promising associations from the meta-analysis, we made use of *in silico* data and imputed genotypes from Harvard and deCODE GWAS data sets together with data from the direct-genotyping Heidelberg-EPIC, ICR, IARC and Toronto replication series.

Harvard. For the Harvard Lung Cancer Susceptibility Study, details of participant recruitment have been described previously[50]. Replication was based on data derived from 1,000 cases and 1,000 controls genotyped using Illumina HumanHap610-Quad arrays. Cases were patients aged >18 years with newly diagnosed, histologically confirmed primary NSCLC. Controls were healthy non–blood related family members and friends of patients with cancer or with cardiothoracic conditions undergoing surgery. The histological classification of lung tumors was performed by two staff pulmonary pathologists at Massachusetts General Hospital according to ICD-O-3: specifically, AD: 8140/3, 8250/3, 8260/3, 8310/3, 8480/3 8560/3; LCC: 8012/3, 8031/3; SQ: 8070/3, 8071/3, 8072/3, 8074/3; other NSCLC: 8010/3, 8020/3, 8021/3, 8032/3, 8230/3. Unqualified samples were excluded if they fit the following quality-control criteria: (i) overall genotype completion rates <95%; (ii) gender discrepancies; (iii) unexpected duplicates or probable relatives (based on a pairwise identity-by-state value of PI_HAT in PLINK >0.185); (iv) heterozygosity rates >6 times the s.d. from the mean; or (v) individuals evaluated to be of non-European ancestry (using HapMap release 23 including the JPT, CEPH, CEU and YRI populations as a reference). Unqualified SNPs were excluded when they fit the following quality-control criteria: (i) SNPs were not mapped on autosomes; (ii) SNPs had a call rate <95% in all GWAS samples; (iii) SNPs had MAF <0.01; or (iv) the genotype distributions of SNPs deviated from those expected by Hardy-Weinberg equilibrium

doi:10.1038/ng.3002

($P < 1.0 \times 10^{-6}$). After applying these prespecified quality controls, genotype data were available for 984 cases and 970 controls.

deCODE. The Icelandic lung cancer study has been described previously[4]. The primary source of information on the Icelandic lung cancer cases is the Icelandic Cancer Registry (ICaR), which covers the entire population of Iceland (http://www.cancerregistry.is/krabbameinsskra/indexen.jsp?id=summary). The sources of data in the ICaR are all pathology and hematology laboratories and all hospital departments and health care facilities in the country. ICaR registration is based on the ICD system and includes information on histology (systemized nomenclature of medicine, SNOMED). ICaR registration also uses the ICD-O system, which takes histology diagnosis into account. Over 94% of diagnoses in the ICaR have histological confirmation. Briefly, according to the ICaR, a total of 4,252 patients were diagnosed with lung cancer from January 1, 1955 to December 31, 2010. Recruitment of both prevalent and incident cases was initiated in 1998, the recruitment is ongoing and DNA samples from lung cancer cases are subjected to whole-genome genotyping as they are collected. The controls used in this study consisted of individuals from other GWAS that were age and sex matched to cases, with no individual disease group accounting for >10% of all controls. Samples were assayed with the Illumina HumanHap300, HumanCNV370, HumanHap610, HumanHap1M, HumanHap660, Omni-1, Omni 2.5 or Omni Express bead chips at deCODE genetics. SNPs were excluded if they had (i) a yield <95%, (ii) MAF < 1% in the population, (iii) deviation from Hardy-Weinberg equilibrium (HWE; $P < 10^{-6}$), (iv) inheritance error rate (>0.001) or (v) if there was a substantial difference in allele frequency between chip types (in which case the SNP was removed from a single chip type if that resolved the difference, but if it did not then the SNP was removed from all chip types). All samples with a call rate of <97% were removed from the analysis. The Icelandic sample set is drawn from the Icelandic population, which is a small homogeneous founder population with almost no detectable population substructure. Thus, there was no need to adjust for such substructure in the association analysis. In addition, the comprehensive Icelandic genealogy database allowed us to exclude individuals not of Icelandic origin from the analysis. SNP genotypes were phased using the method of long-range phasing[51]; for the HumanHap series of chips, 304,937 SNPs were used for long-range phasing, whereas for the Omni series of chips, 564,196 SNPs were used. An initial imputation step was carried out on each chip series separately to create a single harmonized, long-range phased genotype data set consisting of 707,525 SNPs for 95,085 Icelandic individuals. Two sets of genotypes were imputed into this data set with methods previously described[52]: (i) genotypes for about 38 million variants using the 1000 Genomes phase I integrated variant set (v3) as training set and (ii) genotypes for about 34 million variants identified in 2,230 whole genome–sequenced Icelanders. The first set of imputed genotypes was used for replicating the association with variants in the 5p15.33, 9p21 and 12q13.33 regions using IMPUTE (v2.1.1)[53] to perform the case-control analysis. The second set was used when testing the relationship between the p.Lys3326X and c.999del5 genotypes and risk of different cancer types in the Icelandic population using a method that allowed including individuals that had not been chip typed but for whom genotype probabilities were imputed using methods of familial imputation[51].

Heidelberg-EPIC. This study comprised 1,253 Heidelberg-EPIC controls and 1,362 lung cancer cases from the Heidelberg lung cancer study recruited between 1994 and 1998 and between 1996 and 2007, respectively. Details of the Heidelberg-EPIC controls and the Heidelberg lung cancer study have been described previously[54,55]. All subjects were aged 18 years or older, and information on lifestyle risk factors and medical and family history was collected through interviews based on standardized questionnaires. The EPIC Lung and the Heidelberg-EPIC studies were performed independently with no sample overlap with those analyzed as part of the IARC replication series. Histological classification of tumors was obtained from pathology reports, where it was recorded by a staff pulmonary pathologist according to WHO guidelines. Blood samples from patients with malignant lung disease categorized as follows were included: AD, SCLC, NSCLC, LCC, carcinoid, mixed lung tumors or mixed without SCLC. The above-described EPIC Lung and Heidelberg-EPIC studies were performed independently with no sample overlap. Genotypes for SNPs showed no significant departure from HWE, with the exception of rs13314271 in cases.

ICR replication. This study comprised 2,448 cases (1,664 male; mean age at diagnosis 71.8 years, s.d. 6.7 years) with pathologically confirmed lung cancer ascertained through GELCAPS[48] and 2,989 controls (1,469 male; mean age at sampling 60.6 years, s.d. 12.0 years) collected through the National Study of Colorectal Cancer Genetics[56] with no personal history of malignancy. Cases were subclassified into histological subtypes based on ICD coding as described above (in the section detailing the ICR GWAS). Both cases and controls were British residents and had self-reported European ancestry. The genotype distributions of genotypes for each of the SNPs typed in replication showed no significant departure from HWE.

IARC replication. This analysis comprised three studies: (i) EPIC Lung[2,57], a nested case-control study performed within the EPIC (European Prospective Investigation into Cancer and Nutrition) prospective cohort totaling 1,119 lung cancer cases and 2,546 controls (matched one or two to cases for age, sex, center and time of recruitment) selected from 8 of the 10 countries participating in EPIC (Sweden, Netherlands, UK, France, Germany, Spain, Italy and Norway); (ii) the Szczecin case-control study[32], a consecutive series of 849 incident lung cancer cases ascertained from the outpatient oncology clinic in the regional hospital of Szczecin between 2004 and 2007 (the 1,072 controls were individuals without diagnosed cancer or family history of cancer matched to cases by sex, age and region recruited by general medical practitioners); and (iii) Moscow L2, 1,081 newly diagnosed lung cancer cases and 2,119 controls recruited from three hospitals within the Moscow area of Russia between 2007 and 2011. Information on lifestyle risk factors and medical and family history was collected from subjects by interview using a standard questionnaire. Cases were subclassified into histological subtypes based on ICD-O3 coding as described above (in the section detailing the IARC GWAS). The distributions of genotypes for each of the SNPs typed in replication showed no departure from HWE in each country or study series.

Toronto. This study was conducted in the greater Toronto area from 2008 to 2013. Lung cancer cases were recruited at the hospitals in the network of the University of Toronto. Controls were selected randomly from individuals registered in the family medicine clinics databases and were frequency matched with cases on age and sex. All subjects were interviewed, and information on lifestyle risk factors, occupational history and medical and family history was collected using a standard questionnaire. Tumors were centrally reviewed by the reference pathologist (a member of the International Association for the Study of Lung Cancer (IASLC) committee) and a second pathologist in the University Health Network. If the reviews conflicted, a consensus was arrived at after discussion. Coding of histology was based on 2001 WHO/IASLC. After applying standardized quality control procedures and restricting the data to participants with self-reported European ancestry, data and samples were available on 1,084 cases and 966 controls. The genotype distributions of genotypes for each of the SNPs typed in replication showed no significant departure from HWE.

Replication genotyping. Genotyping of rs1519542, rs13314271, rs55731496, rs149423192, rs4592420, rs11571833, rs56084662 and rs17879961 was performed using competitive allele-specific PCR KASPar chemistry (LGC, Hertfordshire, UK; UK replication series), Sequenom (Sequenom, Inc., San Diego, US; Toronto replication and Heidelberg-EPIC replication (rs1519542, rs55731496, rs149423192, rs4592420, rs11571833, rs56084662 and rs17879961)) or TaqMan (Carlsbad, CA; IARC replication series and Heidelberg-EPIC replication (rs13314271)). All primers, probes and conditions used are available on request. Call rates for SNP genotypes were >95% in each of the replication series.

To ensure the quality of genotyping in all assays, at least two negative controls and 1–10% duplicates (showing a concordance of >99%) were genotyped at each center. To exclude technical artifacts in genotyping, at the ICR and IARC we performed cross-platform validation of 96 samples and sequenced a set of 96 randomly selected samples from each case and control series to confirm genotyping accuracy. Assays were found to be performing robustly; concordance was >99%.

Statistical and bioinformatic analyses. Data were imputed for all scans for over 10 million SNPs using data from the 1000 Genomes Project (phase 1

integrated release 3, March 2012) as reference using IMPUTE2 v2.1.1 (ref. 53), MaCH[58] v1.0 or minimac (version 2012.10.3)[59] software (**Supplementary Table 1**). Genotypes were aligned to the positive strand in both imputation and genotyping. Imputation was conducted separately for each scan in which each GWAS data set was pruned to a common set of SNPs between cases and controls before imputation. As previously described, we set thresholds for imputation quality to retain both potential common and rare variants for validation[13,60]. Specifically, poorly imputed SNPs defined by an RSQR < 0.30 with MaCH or an information measure Is < 0.40 with IMPUTE2 were excluded from the analyses. Tests of association between imputed SNPs and lung cancer were performed under a probabilistic dosage model in SNPTEST v2.5 (ref. 61), ProbABEL[62], MaCH2dat v.124 (ref. 58) or the glm function in R. Principle components generated using common SNPs were included in the analysis to limit the effects of cryptic population stratification that might cause inflation of test statistics. The association between each SNP and lung cancer risk was assessed by Cochran-Armitage trend test. The adequacy of the case-control matching and possibility of differential genotyping of cases and controls were formally evaluated using Q-Q plots of test statistics. Meta-analysis was undertaken using inverse-variance approaches. The inflation factor λ was based on the 90% least-significant directly typed SNPs[63]. ORs and associated 95% CIs were calculated by unconditional logistic regression using R (v2.6), Stata v.10 (State College, Texas, US) and PLINK[64] (v1.06) software. Cochran's Q statistic to test for heterogeneity and the I^2 statistic to quantify the proportion of the total variation due to heterogeneity were calculated[65]. I^2 values ≥75% are considered to be characteristic of large heterogeneity[65]. Additionally, analyses stratified by histology, sex, age and smoking status (current, former or never) were performed. All statistical tests were two sided.

The fidelity of imputation as assessed by the concordance between imputed and directly genotyped SNPs was examined in a subset of samples from the UK GWAS, MDACC GWAS, IARC GWAS and NCI GWAS discovery series (**Supplementary Table 2**).

LD metrics were calculated in PLINK using 1000 Genomes data and plotted using SNAP[66]. LD blocks were defined on the basis of HapMap recombination rate (cM/Mb) as defined using the Oxford recombination hotspots and on the basis of the distribution of CIs defined by Gabriel *et al.*[67].

Relationship between genotypes and smoking. To examine the relationship between rs11571833 (BRCA2 p.Lys3326X), rs17879961 (CHEK2 p.Ile157Thr) and rs13314271 (*TP63*) genotype and cigarette consumption (cigarettes per day)[68], we used data on 43,693 Icelandic subjects (including 34,850 chip-typed individuals).

Sequence analysis of *BRCA2* in constitutional DNA. At the ICR, targeted sequencing for the *BRCA2* mutations c.6275delTT and c.4889C>G was performed by Sanger implemented on an ABI3700 analyzer (Applied Biosystems; primer sequences and conditions are available on request). Mutational analysis of the complete coding region of *BRCA2* was based on exome sequencing data generated using Illumina TruSeq capture technology (Illumina, Inc, San Diego, USA). Analysis of Illumina HiSeq2000 (Illumina, Inc, San Diego, USA) sequence data was performed using an in-house pipeline based on the GATK tool kit.

At IARC, Qiagen Generead (SABiosciences/Qiagen Hilde, Germany) was used to amplify the coding region of *BRCA2* in rs11571833 heterozygotes. After library preparation (New England BioLabs, Ipswich, MA, USA), sequencing was performed using an IonTorrent PGM desktop sequencer (Life Technologies, Guilford, San Francisco, CA). Genotypes were called using Ionsuite software. Sequence changes were referenced to the Leiden Open Variation Database (LOVD2) and the BReast CAncer IARC database.

Analysis of TCGA data. The exomes of 243 individuals with lung SQ and 338 individuals with lung AD in TCGA (Project Number #3230) were analyzed at IARC using an in-house pipeline based on the GATK tool set. Variant calls were annotated using ANNOVAR, making use of use the National Heart, Lung, and Blood Institute's Exome Sequencing Project and 1000 Genomes data.

Copy number variation. Copy number variation was assessed from Human SNP Array 6.0 data. We retrieved level 3 TCGA data comprising normalized $\log_2$ ratios of the fluorescence intensities between the target sample and a reference sample. We included only tumor-normal paired data in our analysis. We considered a $\log_2$ ratio ≤0.5 as reflecting loss and a $\log_2$ ratio >0.5 as reflecting gain. Annotation was performed by adding the genes contained in each of the remaining segments using EnsEMBL databases.

38. Omenn, G.S. *et al.* The β-carotene and retinol efficacy trial (CARET) for chemoprevention of lung cancer in high risk populations: smokers and asbestos-exposed workers. *Cancer Res.* **54** (suppl. 7), 2038s–2043s (1994).
39. Scélo, G. *et al.* Occupational exposure to vinyl chloride, acrylonitrile and styrene and lung cancer risk (Europe). *Cancer Causes Control* **15**, 445–452 (2004).
40. Feyler, A. *et al.* Point: myeloperoxidase –463G→A polymorphism and lung cancer risk. *Cancer Epidemiol. Biomarkers Prev.* **11**, 1550–1554 (2002).
41. Nelis, M. *et al.* Genetic structure of Europeans: a view from the North-East. *PLoS ONE* **4**, e5472 (2009).
42. Välk, K. *et al.* Gene expression profiles of non-small cell lung cancer: survival prediction and new biomarkers. *Oncology* **79**, 283–292 (2010).
43. Holmen, J. *et al.* The Nord-Trondelag Health Study 1995–97 (HUNT2): objectives, contents, methods and participation. *Norsk Epidemiologi* **13**, 1932 (2003).
44. Landi, M.T. *et al.* Environment And Genetics in Lung cancer Etiology (EAGLE) study: an integrative population-based case-control study of lung cancer. *BMC Public Health* **8**, 203 (2008).
45. ATBC Cancer Prevention Study Group. The α-tocopherol, β-carotene lung cancer prevention study: design, methods, participant characteristics, and compliance. *Ann. Epidemiol.* **4**, 1–10 (1994).
46. Hayes, R.B. *et al.* Methods for etiologic and early marker investigations in the PLCO trial. *Mutat. Res.* **592**, 147–154 (2005).
47. Calle, E.E. *et al.* The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. *Cancer* **94**, 2490–2501 (2002).
48. Eisen, T., Matakidou, A. & Houlston, R. Identification of low penetrance alleles for lung cancer: the GEnetic Lung CAncer Predisposition Study (GELCAPS). *BMC Cancer* **8**, 244 (2008).
49. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
50. Su, L. *et al.* Genotypes and haplotypes of matrix metalloproteinase 1, 3 and 12 genes and the risk of lung cancer. *Carcinogenesis* **27**, 1024–1029 (2006).
51. Kong, A. *et al.* Parental origin of sequence variants associated with complex diseases. *Nature* **462**, 868–874 (2009).
52. Styrkarsdottir, U. *et al.* Nonsense mutation in the *LGR4* gene is associated with several human diseases and other traits. *Nature* **497**, 517–520 (2013).
53. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
54. Boeing, H., Wahrendorf, J. & Becker, N. EPIC-Germany—a source for studies into diet and risk of chronic diseases. European Investigation into Cancer and Nutrition. *Ann. Nutr. Metab.* **43**, 195–204 (1999).
55. Dally, H. *et al.* The *CYP3A4*1B* allele increases risk for small cell lung cancer: effect of gender and smoking dose. *Pharmacogenetics* **13**, 607–618 (2003).
56. Penegar, S. *et al.* National study of colorectal cancer genetics. *Br. J. Cancer* **97**, 1305–1309 (2007).
57. Timofeeva, M.N. *et al.* Genetic polymorphisms in 15q25 and 19q13 loci, cotinine levels, and risk of lung cancer in EPIC. *Cancer Epidemiol. Biomarkers Prev.* **20**, 2250–2261 (2011).
58. Li, Y., Willer, C.J., Ding, J., Scheet, P. & Abecasis, G.R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.* **34**, 816–834 (2010).
59. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* **44**, 955–959 (2012).
60. Zeggini, E. *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40**, 638–645 (2008).
61. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* **11**, 499–511 (2010).
62. Aulchenko, Y.S., Struchalin, M.V. & van Duijn, C.M. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* **11**, 134 (2010).
63. Clayton, D.G. *et al.* Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat. Genet.* **37**, 1243–1246 (2005).
64. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
65. Higgins, J.P., Thompson, S.G., Deeks, J.J. & Altman, D.G. Measuring inconsistency in meta-analyses. *Br. Med. J.* **327**, 557–560 (2003).
66. Johnson, A.D. *et al.* SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24**, 2938–2939 (2008).
67. Gabriel, S.B. *et al.* The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).
68. Thorgeirsson, T.E. *et al.* Sequence variants at *CHRNB3-CHRNA6* and *CYP2A6* affect smoking behavior. *Nat. Genet.* **42**, 448–453 (2010).

doi:10.1038/ng.3002

Oncogenic and drug-sensitive *NTRK1* rearrangements in lung cancer

Aria Vaishnavi[1,13], Marzia Capelletti[2,13], Anh T Le[1], Severine Kako[3], Mohit Butaney[2], Dalia Ercan[2], Sakshi Mahale[3], Kurtis D Davies[1], Dara L Aisner[3,4], Amanda B Pilling[1], Eamon M Berge[1], Jhingook Kim[5], Hidefumi Sasaki[6], Seung-il Park[7], Gregory Kryukov[8], Levi A Garraway[8,9], Peter S Hammerman[2], Julia Haas[10], Steven W Andrews[10], Doron Lipson[11], Philip J Stephens[11], Vince A Miller[11], Marileila Varella-Garcia[1,3], Pasi A Jänne[2,12,14] & Robert C Doebele[1,3,14]

We identified new gene fusions in patients with lung cancer harboring the kinase domain of the *NTRK1* gene that encodes the high-affinity nerve growth factor receptor (TRKA protein). Both the *MPRIP-NTRK1* and *CD74-NTRK1* fusions lead to constitutive TRKA kinase activity and are oncogenic. Treatment of cells expressing *NTRK1* fusions with inhibitors of TRKA kinase activity inhibited autophosphorylation of TRKA and cell growth. Tumor samples from 3 of 91 patients with lung cancer (3.3%) without known oncogenic alterations assayed by next-generation sequencing or fluorescence *in situ* hybridization demonstrated evidence of *NTRK1* gene fusions.

Treatment with orally active kinase inhibitors crizotinib and erlotinib or gefitinib is superior to standard chemotherapy in patients with lung cancers that have *ALK* fusions or *EGFR* mutations, respectively[1,2]. Additional oncogenes involving fusions of *ROS1*, *RET*, *FGFR1*, *FGFR2* and *FGFR3* have been identified in lung cancers and demonstrate great potential for therapeutic intervention[3–9]. These oncogenes also occur in several other common malignancies, expanding the potential relevance of this therapeutic approach[9–12].

We performed a targeted next-generation DNA sequencing (NGS) assay on tumor samples from 36 patients with lung adenocarcinoma whose tumors did not contain known genetic alterations identifiable using standard clinical assays[10] (**Supplementary Table 1**).

We detected evidence of an in-frame gene fusion event in 2 of 36 patients, both of whom were female never-smokers with lung adenocarcinoma, involving the kinase domain of the *NTRK1* gene, which encodes the TRKA receptor tyrosine kinase but no other potentially oncogenic alterations (**Fig. 1**, **Supplementary Fig. 1** and **Supplementary Table 1**). We confirmed the exon junctions and mRNA expression by RT-PCR and cloning of the entire cDNAs (**Supplementary Figs. 2–4**). In the first case, the 5′ end of the myosin phosphatase–Rho-interacting protein gene (*MPRIP*) was joined with the 3′ end of *NTRK1*. *MPRIP* is involved in actin cytoskeleton regulation and has been implicated in a gene fusion in small-cell lung cancer, putatively causing early termination of *TP53* (ref. 13). We detected expression of the fusion protein, RIP-TRKA (encoded by *MPRIP-NTRK1*), in 293T cells with exogenous expression of the *MPRIP-NTRK1* cDNA, a malignant pleural effusion sample and early-passage lung cancer–derived cells (CUTO-3) derived from the same patient growing in culture (**Supplementary Fig. 4**). CUTO-3 cells demonstrated autophosphorylation of this previously uncharacterized protein at critical TRKA tyrosine residues[14]. MPRIP harbors three coiled-coil domains (CCDs), one of which mediates interaction with myosin phosphatase[15]. Gene partners of ALK, ROS1, and RET fusions often contain CCDs, which are known to mediate dimerization and consequent activation of the kinase domain; thus, it is likely that the CCDs contained within *MPRIP-NTRK1* perform a similar function[4,16] (**Supplementary Fig. 5**). The second case harbored a *CD74-NTRK1* gene fusion. *CD74*, which encodes the major histocompatibility complex (MHC) class II invariant chain, is a known activating fusion partner of *ROS1*, and the CD74-TRKA protein is predicted to be localized in the plasma membrane[3,17–19] (**Supplementary Fig. 5**).

We developed a fluorescence *in situ* hybridization (FISH) assay to detect chromosomal rearrangements within the *NTRK1* gene (**Supplementary Fig. 6a**). Hybridization of these probes showed clear separation of the 5′ and 3′ probes in the tumor cells in the patient samples containing the *MPRIP-NTRK1* or *CD74-NTRK1* gene fusions, but not in the nontumor cells from those samples or a control sample (**Fig. 1b** and **Supplementary Fig. 6b**). Fusions between *NTRK1* and *TPM3*, *TFG* or *TPR* have previously been identified in colorectal and thyroid cancers[11,20]. Although *TPM3* (1q22-23) lies in close proximity to *NTRK1* (1q21-22), FISH detected a separation in signals in the KM12 colorectal cancer cell line that harbors a *TPM3-NTRK1* fusion (**Supplementary Figs. 6c** and **7**). Using this FISH assay, 56 additional lung adenocarcinoma samples without detectable oncogenic alterations were screened for *NTRK1* rearrangements, and one additional positive case was identified (**Supplementary Table 2** and **Supplementary Fig. 6d**). Quantitative

[1]Department of Medicine, Division of Medical Oncology, University of Colorado School of Medicine, Aurora, Colorado, USA. [2]Lowe Center for Thoracic Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [3]University of Colorado Cancer Center, Aurora, Colorado, USA. [4]Department of Pathology, University of Colorado School of Medicine, Aurora, Colorado, USA. [5]Department of Thoracic Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea. [6]Department of Oncology, Immunology and Surgery, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan. [7]Department of Thoracic Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea. [8]Broad Institute, Cambridge, Massachusetts, USA. [9]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [10]Array BioPharma, Boulder, Colorado, USA. [11]Foundation Medicine, Boston, Massachusetts, USA. [12]Belfer Institute for Applied Cancer Science, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [13]These authors contributed equally to this work. [14]These authors jointly directed this work. Correspondence should be addressed to R.C.D. (robert.doebele@ucdenver.edu) or P.A.J. (pasi_janne@dfci.harvard.edu).

Received 28 March; accepted 15 August; published online 27 October 2013; doi:10.1038/nm.3352

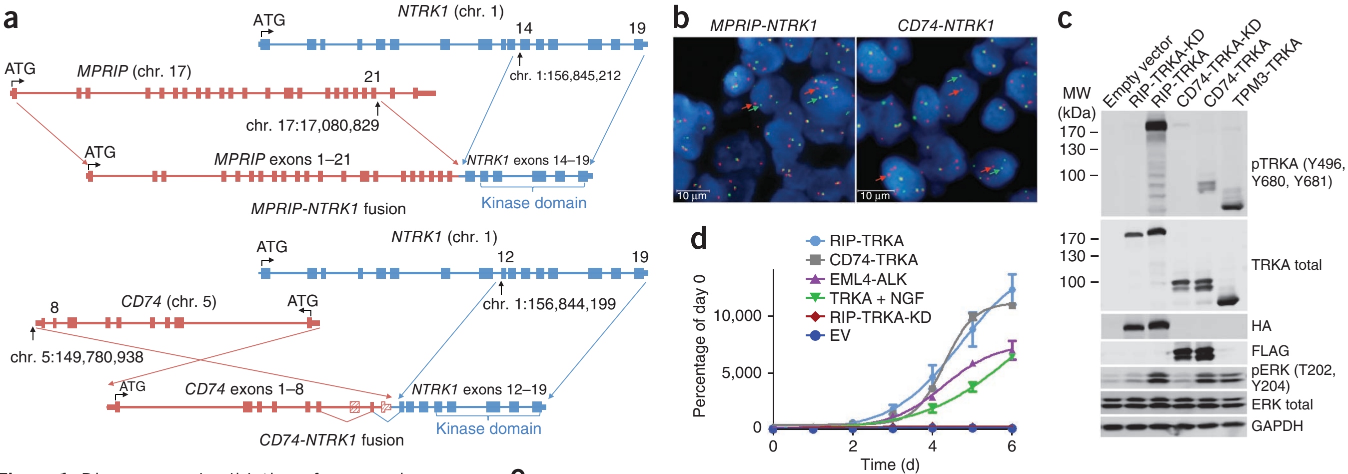
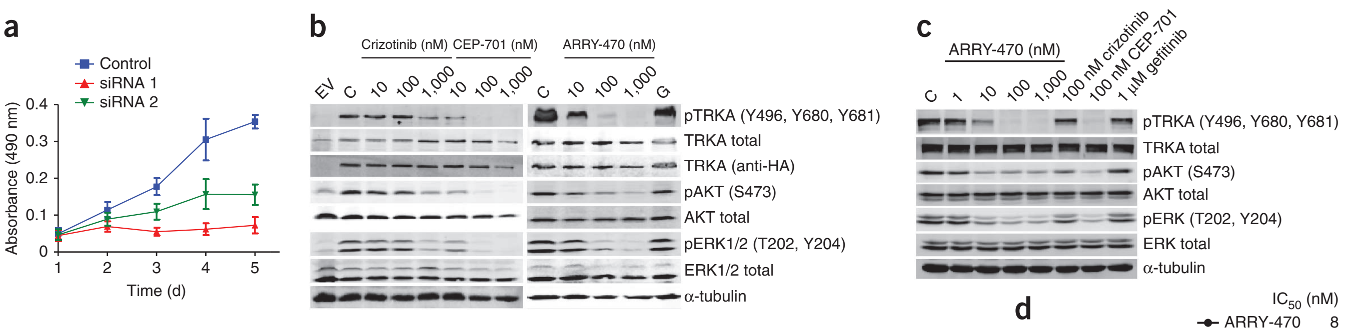


Figure 1 Discovery and validation of oncogenic *NTRK1* gene fusions in lung cancer samples. (**a**) Schematic of genomic rearrangement from tumor samples harboring *MPRIP-NTRK1* and *CD74-NTRK1* using the FoundationOne next-generation sequencing assay, including chromosomal breakpoints for each gene rearrangement. (**b**) Break-apart FISH analysis of *MPRIP-NTRK1* and *CD74-NTRK1* tumor samples showing clear separation of green (5′) and red (3′) signals corresponding to the *NTRK1* gene. (**c**) TRKA (*NTRK1*) fusions are autophosphorylated and activate key downstream signaling pathways. Representative immunoblot analyses ($n = 3$) of cell lysates from 293T cells expressing RIP-TRKA and CD74-TRKA, but not their kinase-dead (KD) variants, display phosphorylation of critical tyrosine residues (Y496, Y680 and Y681) in TRKA (pTRKA) and activation of ERK (phosphorylation of T202 and Y204, pERK). TPM3-TRKA was expressed in 293T cells as a positive control. MW, molecular weight; HA, hemagglutinin. (**d**) *NTRK1* fusions support cellular proliferation. 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) assay of Ba/F3 demonstrated that cells expressing RIP-TRKA, CD74-TRKA, EML4-ALK or full-length TRKA supplemented with NGF proliferated in the absence of IL-3, whereas Ba/F3 cells expressing empty vector (EV) or the kinase-dead variant of RIP-TRKA did not proliferate ($n = 3$). Data are expressed as mean ± s.e.m. (**e**) *MPRIP-NTRK1* or *CD74-NTRK1* gene fusions induce tumorigenesis. NIH3T3 cells expressing RIP-TRKA, RIP-TRKA-KD, CD74-TRKA, EML4-ALK or empty vector were injected into the flanks of nude mice and observed for tumor growth. Representative pictures taken at day 12 following injection are shown. The number of tumors induced in the injected animals are shown in parentheses (five mice injected per cell line, $n = 1$).

PCR demonstrated high *NTRK1* kinase domain expression only in the tumors with the known *NTRK1* rearrangements or in the KM12 cell line (**Supplementary Fig. 8**). Analysis of transcriptome data from The Cancer Genome Atlas of 230 lung adenocarcinomas failed to detect evidence of *NTRK1* fusions (data not shown). A recent transcriptome study of 87 lung adenocarcinoma tumor samples, which



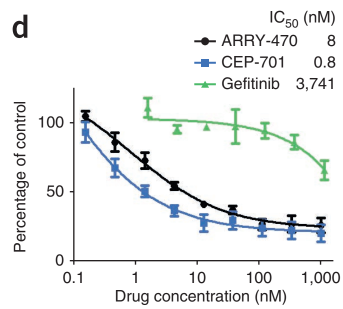


Figure 2 Drug treatment inhibits activation of TRKA, downstream signaling and proliferation in cells expressing *NTRK1* fusions. (**a**) RNAi knockdown of *NTRK1* inhibits cell proliferation in a cell line harboring *TPM3-NTRK1*. KM12 cells were analyzed by MTS proliferation assay 1–5 d after transfection with two different *NTRK1*-specific siRNAs (siRNA 1 and siRNA 2) or scrambled siRNA (control) ($n = 3$). Data are expressed as mean ± s.e.m. (**b**) Ba/F3 cells expressing *MPRIP-NTRK1* (RIP-TRKA) or EV were lysed after 5 h of treatment with the indicated doses of drugs (G, gefitinib 1,000 nM) or DMSO control (C). Phosphorylation of TRKA (Y496, Y680 and Y681) or ERK (T202 and Y204) was assessed by antibodies specific to the indicated tyrosine or threonine residues. (**c**) CUTO-3 lung cancer cells harboring the *MPRIP-NTRK1* gene fusion were treated with the indicated doses and drugs and subjected to immunoblot analysis. (**d**) Ba/F3 cells expressing the *MPRIP-NTRK1* fusion demonstrate inhibition of proliferation as measured by MTS assay using the pan-TRK inhibitor, ARRY-470, and the multikinase inhibitor, CEP-701, but not the EGFR inhibitor, gefitinib ($n = 5$). Data are expressed as mean ± s.e.m.

identified a set of previously uncharacterized fusions, did not include oncogenic fusions involving *NTRK1* (J.-S. Seo, Seoul National University College of Medicine, personal communication)[21].

To formally prove that these fusion proteins are oncogenic, we expressed *MPRIP-NTRK1* and *CD74-NTRK1* cDNA constructs in three other noncancer cell lines commonly used to assess oncogenicity—293T cells, NIH3T3 fibroblasts and Ba/F3 cells—and assessed them for oncogenic traits expected in these cells. We observed expression of the appropriate-sized chimeric proteins and TRKA autophosphorylation, as seen in the CUTO-3 cells[14] (**Figs. 1c** and **2b** and **Supplementary Figs. 4**, **9** and **14**). Introduction of a *NTRK1* kinase-dead mutation did not result in TRKA autophosphorylation or in increased phosphorylation of extracellular signal–related kinases 1 and 2 (ERK1/2) (**Figs. 1c** and **Supplementary Fig. 9**). *MPRIP-NTRK1* and *CD74-NTRK1*, but not their kinase-dead counterparts, induced interleukin-3 (IL-3)-independent proliferation of Ba/F3 cells (**Fig. 1d**). Similarly, *MPRIP-NTRK1* and *CD74-NTRK1* supported anchorage-independent growth of NIH3T3 cells and formed tumors in nude mice, and *CD74-NTRK1* induced a light-refractory appearance of NIH3T3 cells (**Fig. 1e** and **Supplementary Figs. 10** and **11**). Knockdown of *TPM3-NTRK1* in KM12 cells reduced proliferation, further supporting the role of *NTRK1* fusions as oncogenes (**Fig. 2a** and **Supplementary Fig. 12**).

Given the prior success of treatment with kinase inhibitors in patients with cancers harboring *ALK* and *ROS1* fusions, we asked whether *NTRK1* fusions might provide a similar target in patients with lung cancer or other malignancies. ARRY-470 is a selective kinase inhibitor with nanomolar activity against TRKA, TRKB and TRKC but no other notable kinase inhibition below 1,000 nM (**Supplementary Fig. 13** and **Supplementary Table 3**). Lestaurtinib (CEP-701) and crizotinib also have activity against TRKA in addition to other kinases[22,23]. Treatment of cells expressing *MPRIP-NTRK1* or *CD74-NTRK1* with ARRY-470, CEP-701 and, to a lesser extent, crizotinib inhibited autophosphorylation of RIP-TRKA and CD74-TRKA (**Fig. 2b** and **Supplementary Figs. 9** and **14a**). Activation of the mitogen-activated protein kinase (MAPK) and AKT kinase pathways was also inhibited in Ba/F3 cells expressing exogenous *MPRIP-NTRK1* or *CD74-NTRK1* (**Fig. 2b** and **Supplementary Fig. 14a**). Phosphorylation of endogenously expressed RIP-TRKA in CUTO-3 and tropomyosin 3 (TPM3)-TRKA in KM12 cells was similarly inhibited by all three drugs (**Fig. 2c** and **Supplementary Fig. 15a**).

Inhibition of proliferation of Ba/F3 cells expressing *NTRK1* gene fusions was greatest with CEP-701 and ARRY-470 (**Fig. 2d** and **Supplementary Fig. 14b**). Crizotinib was a less potent inhibitor, although in a range similar to that seen for inhibition of echinoderm microtubule-associated protein-like 4–anaplastic lymphoma kinase (EML4-ALK) or syndecan-4–c-ros oncogene 1, receptor tyrosine kinase (SDC4-ROS1) (ref. 3) (**Supplementary Fig. 16**). The less potent effects of crizotinib on cell proliferation are consistent with decreased inhibition of phosphorylated TRKA (pTRKA) and downstream phosphorylated ERK1/2 (pERK1/2) (**Fig. 2b** and **Supplementary Fig. 14a**). All three drugs also inhibited colony formation of NIH3T3 cells expressing *NTRK1* fusions in soft agar (**Supplementary Fig. 10**). KM12 cells were similarly sensitive to ARRY-470 and CEP-701, but less so to crizotinib (**Supplementary Fig. 15b**). ARRY-470 did not inhibit proliferation of Ba/F3 cells expressing other oncogene targets (epidermal growth factor receptor (EGFR), ALK or ROS1) or of lung and colorectal cell lines that do not harbor an *NTRK1* fusion (**Supplementary Fig. 17**). All three drugs induced cell-cycle arrest in G1 and apoptosis of KM12 cells (**Supplementary Fig. 18**).

The index patient (*MPRIP-NTRK1*) had no standard therapies and no clinical trials of potentially effective TRKA inhibitors available; therefore, the patient consented to treatment with crizotinib (250 mg twice daily) outside of a clinical trial. The patient showed a minor radiographic response with a decrease in serum levels of tumor marker CA125 but experienced disease progression after ~3 months (**Supplementary Fig. 19**). This modest clinical activity of crizotinib is consistent with the *in vitro* activity that we observed and could be caused by non–TRKA kinase effects.

We have identified new recurrent oncogenic *NTRK1* fusions in a subset of patients (3 out of 91, 3.3%) with lung adenocarcinoma that did not contain other common oncogenic alterations. Our study further highlights the utility of targeted NGS to discover drug-sensitive genetic alterations in patients with lung cancer. Based on our data, clinical studies of selective TRKA inhibitors in *NTRK1*–rearranged non–small-cell lung cancer are warranted.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. cDNA sequences are available in GenBank under accession codes KF724384 and KF724385.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was supported by the Colorado Bioscience Discovery Evaluation Grant Program, the National Institutes of Health (NIH) NCATS Colorado Clinical and Translational Sciences Institute Grant UL1 TR000154 (contents are the authors' sole responsibility and do not necessarily represent official NIH views) and the Boettcher Foundation's Webb-Waring Biomedical Research Program (R.C.D.), as well as the Dana-Farber/Harvard Cancer Center Lung Cancer Specialized Programs of Research Excellence (SPORE) P50 CA090578 (P.A.J.), the Cammarata Family Foundation Research Fund (M.C. and P.A.J.), the Nirenberg Fellowship at the Dana-Farber Cancer Institute (M.C. and P.A.J.) and the University of Colorado Lung Cancer SPORE grant P50CA058187 and the US National Cancer Institute Cancer Center Support Grant P30CA46934 (M.V.-G.). Ba/F3 cells were a gift from D. Theodorescu (University of Colorado, Anschutz Medical Campus). The cell lines KM12, HCT116, HCT15, HT29 and SW837 were gifts from S.G. Eckhardt (University of Colorado, Anschutz Medical Campus). The cell lines H3122, H1650, H1299 and HCC78 were gifts from J.D. Minna (University of Texas Southwestern). Aythymic nude mice were a gift from J. DeGregori (University of Colorado, Anschutz Medical Campus).

AUTHOR CONTRIBUTIONS

R.C.D., P.A.J. and V.A.M. conceived of the research idea. R.C.D. and P.A.J. designed experiments, took responsibility for the oversight of the project and wrote the manuscript. A.V. and D.E. designed and performed immunoblotting, generated Ba/F3 and NIH3T3 cells expressing exogenous *NTRK1* fusion constructs, performed flow cytometry and MTS assays and contributed to writing of the manuscript. D.L. and P.J.S. designed and performed NGS assays. M.V.-G. designed FISH probes and interpreted FISH experiments. S.K. and S.M. performed and analyzed FISH experiments. A.T.L. and M.C. performed cloning of RT-PCR fusion constructs. D.L.A. analyzed tumor samples and performed immunohistochemistry. K.D.D. performed *in vitro* analyses and contributed to interpretation of the data. A.B.P. performed *in vivo* experiments and contributed to interpretation of data. P.S.H., L.A.G. and G.K. performed bioinformatics analyses. J.H. and S.W.A. designed and profiled ARRY-470. E.M.B. and M.B. collected clinical data. J.K., H.S. and S.P. provided subject specimens. All authors contributed to revision of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Mok, T.S. *et al. N. Engl. J. Med.* **361**, 947–957 (2009).
2. Shaw, A.T. *et al. N. Engl. J. Med.* **368**, 2385–2394 (2013).
3. Davies, K.D. *et al. Clin. Cancer Res.* **18**, 4570–4579 (2012).
4. Takeuchi, K. *et al. Nat. Med.* **18**, 378–381 (2012).
5. Ju, Y.S. *et al. Genome Res.* **22**, 436–445 (2012).
6. Kohno, T. *et al. Nat. Med.* **18**, 375–377 (2012).
7. Drilon, A. *et al. Cancer Discov.* **3**, 630–635 (2013).
8. Majewski, I.J. *et al. J. Pathol.* **230**, 270–276 (2013).
9. Wu, Y.M. *et al. Cancer Discov.* **3**, 636–647 (2013).
10. Lipson, D. *et al. Nat. Med.* **18**, 382–384 (2012).
11. Alberti, L., Carniti, C., Miranda, C., Roccato, E. & Pierotti, M.A. *J. Cell Physiol.* **195**, 168–186 (2003).
12. Davies, K.D. & Doebele, R.C. *Clin. Cancer Res.* **19**, 4040–4045 (2013).
13. Peifer, M. *et al. Nat. Genet.* **44**, 1104–1110 (2012).
14. Stephens, R.M. *et al. Neuron* **12**, 691–705 (1994).
15. Surks, H.K., Richards, C.T. & Mendelsohn, M.E. *J. Biol. Chem.* **278**, 51484–51493 (2003).
16. Soda, M. *et al. Nature* **448**, 561–566 (2007).
17. Busch, R., Doebele, R.C., Patil, N.S., Pashine, A. & Mellins, E.D. *Curr. Opin. Immunol.* **12**, 99–106 (2000).
18. Rikova, K. *et al. Cell* **131**, 1190–1203 (2007).
19. Awad, M.M. *et al. N. Engl. J. Med.* **369**, 1173 (2013).
20. Martin-Zanca, D., Hughes, S.H. & Barbacid, M. *Nature* **319**, 743–748 (1986).
21. Seo, J.S. *et al. Genome Res.* **22**, 2109–2119 (2012).
22. George, D.J. *et al. Cancer Res.* **59**, 2395–2401 (1999).
23. Cui, J.J. *et al. J. Med. Chem.* **54**, 6342–6363 (2011).

ONLINE METHODS

Patient tumor samples. Colorado Multiple Institutional Review Board (IRB) or Dana-Farber Cancer Institute IRB approval was obtained for all patients in this study. All patients provided written informed consent. FoundationOne testing and FISH analyses were performed in CLIA-certified laboratories. The index patient who underwent treatment with crizotinib consented to treatment outside of a clinical trial.

Next-generation DNA sequencing. DNA was extracted from 40 μm of formalin-fixed paraffin-embedded (FFPE) or frozen tissue using the Maxwell 16 FFPE Plus LEV DNA Purification kit (Promega) and quantified using the PicoGreen fluorescence assay (Invitrogen). Library construction was performed as previously described using 50–200 ng of DNA sheared by sonication to ~100–400 base pairs before end repair, dA addition and ligation of indexed Illumina sequencing adaptors[24]. Enrichment of target sequences (3,320 exons of 182 cancer-related genes and 37 introns from 14 genes recurrently rearranged in cancer representing ~1.1 Mb of the human genome in total) was achieved by solution-based hybrid capture with a custom Agilent SureSelect biotinylated RNA baitset[24]. The libraries were sequenced on an Illumina HiSeq 2,000 platform using 49 × 49 paired-end reads. Sequence data from genomic DNA was mapped to the reference human genome (hg19) using the Burrows-Wheeler Aligner and were processed using the publicly available SAMtools, Picard and Genome Analysis Toolkit[25,26]. Genomic rearrangements were detected by clustering chimeric reads mapped to targeted introns.

RNA extraction from formalin-fixed paraffin-embedded and frozen tissues. RNA was isolated from FFPE or frozen tumor samples as described previously[3]. Briefly, FFPE samples were processed using the RecoverAll Total Nucleic Acid Isolation Kit (Ambion) following deparaffinization in xylene and washed with 100% ethanol before Protease K digestion. Extraction of RNA from frozen tissue was accomplished using TriReagent (Ambion). Alternatively, tumors from patients with non–small cell lung cancer obtained at surgery were snap-frozen in liquid nitrogen, embedded in optimal cutting temperature (OCT) medium and sectioned. RNA was prepared using Trizol (Invitrogen) followed by RNeasy MinElute cleanup kit (Qiagen).

RT-PCR and sequencing of *MPRIP-NTRK1* and *CD74-NTRK1*. RT-PCR of *MPRIP-NTRK1* was carried out using the SuperScript III First-Strand Synthesis System (SSIII RT) from Invitrogen with a *NTRK1* primer located in exon 15 ('*NTRK1* Y490R1') followed by PCR using the same reverse primer, *NTRK1* Y490R1, and a primer to *MPRIP* located in its third coil-coiled domain ('*MPRIP* CC3F1'). PCR products were resolved on an agarose gel, and the fragments were excised and treated with ExoSapIT (Affymetrix) before being sequenced by the University of Colorado Cancer Center DNA Sequencing and Analysis Core using the BigDye Terminator Cycle Sequencing Ready Reaction kit version 1.1 (Applied Biosystems) using the same forward and reverse primer as in the RT-PCR reaction. For *CD74-NTRK1*, reverse transcription was carried out using the QuanTitec Reverse Transcription Kit (Qiagen). PCR of the resulting cDNA was performed using the primers 'CD74 Exon 3 FOR' and 'NTRK1 Exon 15 REV'. Primers used for RT-PCR and sequencing are available in **Supplementary Table 4**. The reference sequences used for exon alignment are National Center for Biotechnology Information (NCBI) reference sequences NM_002529.3 (*NTRK1*), NM_015134.3 (*MPRIP*) and NM_001025159.2 (*CD74*).

Cloning full-length *MPRIP-NTRK1, CD74-NTRK1* and *TPM3-NTRK.* cDNA was generated from each patient tumor sample using the following procedures. For the *MPRIP-NTRK1* construct, cDNA was transcribed using the SSIII RT kit from Invitrogen along with a primer located at the end of *NTRK1* (NTRK1stopR2). This cDNA was used to amplify two separate overlapping fragments that were used to generate full-length *MPRIP-NTRK1* by overlap-extension PCR using the two fragments alone for ten cycles and then adding the MPRIPStart and NTRK1stopR1 primers for an additional 30 cycles of PCR amplification. The resulting 4-kb PCR product was gel-isolated and confirmed by Sanger sequencing. A 3′ hemagglutinin (HA) tag was added to *MPRIP-NTRK1* using PCR amplification with primers harboring the HA-encoding sequence. The amplified product was subsequently cloned into the pCDH-CMV-MSC1-EF1-Puro lentiviral plasmid (System Biosciences). Full-length *TPM3-NTRK1* was amplified from KM12 cDNA using TPM3Start RI and NTRKStopNotI primers and cloned into the lentiviral plasmid as described above. The National Center for Biotechnology Information (NCBI) reference sequence used for *TPM3* is NM_153649.3. For the *CD74-NTRK1* construct, cDNA was transcribed with Quantiscript Reverse Transcriptase (Qiagen). Full-length *CD74-NTRK1* was amplified using the primers 'CD74 FOR' and 'NTRK1 REV' using AccuPrime Taq DNA Polymerase (Invitrogen) and cloned into the pDNR-Dual vector (BD Biosciences) and recombined into JP1520 retroviral vector as previously described[27]. The full-length cDNA of each gene was confirmed by sequencing. Primers used for cloning are available in **Supplementary Table 4**.

Quantitative PCR of *NTRK1*. Relative-quantification PCR (RQ-PCR) assay of the *NTRK1* tyrosine kinase domain (Hs01021011_m1; Applied Biosystems) was used to evaluate its level of mRNA expression. The relative quantification method ($\Delta\Delta C_T$) in the StepOnePlus Real-Time PCR system (Applied Biosystems) was used with β-glucuronidase (GUSB) (Applied Biosystems) as an endogenous control. All samples were evaluated in triplicate.

RNA sequencing. Paired-end RNA sequencing was performed as previously described[28]. RNA FASTQ files were aligned and splice junctions mapped using TopHat[29] and analyzed for fusion reads using the Broad Institute Cancer Genome Analysis Tools Suite (http://www.broadinstitute.org/cancer/cga/ and http://www.broadinstitute.org/cancer/software/genepattern/modules/RNA-seq/).

Cell lines and reagents. NIH3T3, HEK-293T and A549 cells were purchased from ATCC. NIH3T3, HEK-293T and Ba/F3 were previously described[30]. The lung cancer cell lines A549, H3122, H1650, H1299 and HCC78 were previously described[30–33]. The colorectal cancer cell lines KM12, HCT116, HCT15, HT29 and SW837 were previously described[34]. Ba/F3 cells expressing the mutant *EGFR* allele, E746_A750del, were previously described[27]. The lymphoblastoid cell line GM09948 (Coriell Cell Repository) was used for genomic mapping in FISH studies.

All cancer cell lines were maintained in RPMI medium with 10% FBS. NIH3T3 and Ba/F3 cells transduced with full-length *NTRK1* were supplemented with 100 ng/ml and 200 ng/ml β nerve growth factor (β-NGF) (R&D Systems), respectively. Crizotinib and gefitinib were purchased from Selleck Chemicals, CEP-701 from Sigma Aldrich or Santa Cruz Biotechnology and K252a from Tocris, and ARRY-470 was supplied by Array BioPharma. Antibodies to total AKT (clone 40D4, 1:2,000), AKT pSer473 (clone D9E, 1:5,000), total ERK (clone L34F12, 1:2,000), ERK pThr202/Tyr204 (clone D13.14.4E, 1:5,000), total STAT3 (clone 124H6, 1:5,000), STAT3 pY705 (clone D3A7, 1:2,000), PARP (clone 46D11, 1:5,000) and TRK pY490 (cat. #9141, corresponding Y496 in TRKA, 1:1,000) and pY674/675 (clone C50F3, corresponding to Y680 and Y681 in TRKA, 1:1,000) were purchased from Cell Signaling Technology. Total TRKA (clone C-14, 1:1,000), glyceraldehyde 3-phosphate dehydrogenase (clone MAB374. 1:5,000) and α-tubulin (clone TU-02, 1:5,000) were purchased from Santa Cruz Biotechnology.

Derivation and propagation of CUTO-3 cells. The patient gave written informed consent for the derivation of a cancer cell line. Pleural fluid from the index patient harboring the *MPRIP-NTRK1* gene was collected and mononuclear cells were isolated by centrifugation through a ficoll gradient (Thermo Scientific). Cells were seeded onto a 25-cm flask and cultured in serum-free ACL4 medium to inhibit outgrowth of normal stromal cells[35]. Once the tumor cell became the predominate cell type in the culture flask, the culture was subjected to differential trypsinization in order to dislodge the remaining minor population of stromal cells. After this enrichment process, tumors cells were cultured in ACL4 medium supplemented with 5% heat-inactivated FBS and routinely passaged using this medium. The CUTO-3 cells were later adapted to grow in RPMI 1640 with 10% FBS for ease of culturing and experimentation.

Lentivirus or retrovirus production and cell transduction. *MPRIP-NTRK1* or the kinase-dead variant was introduced into cells by lentivirus as previously described[30]. NIH3T3 cells transduced with lentivirus were cultured in DMEM medium with 5% FCS and 0.75 μg/ml puromycin. Ba/F3 cells transduced with lentivirus were cultured as above with 2 μg/ml puromycin and with or without

1 ng/ml IL-3 (R&D Systems). Alternatively, *CD74-NTRK1* was introduced into cells using retrovirus as previously described[27]. Polyclonal cell lines were established by puromycin selection. Cell proliferation and growth were performed as previously described[27,36].

Mouse xenograft studies. NIH3T3 cells (1×10^6) harboring the indicated expression vectors were resuspended in Matrigel (BD Biosciences) and injected subcutaneously into athymic nude mice (gift from J. DeGregori). Mice were monitored three times weekly for tumor formation and sacrificed when tumors reached approximately 2 cm × 2 cm. Approval for the use of animals in this study was granted by the University of Colorado Institutional Animal Care and Use Committee.

Immunoblotting. Immunoblotting was performed as previously described[30]. Briefly, cells were lysed in RIPA buffer with Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific) and diluted in loading buffer (LI-COR Biosciences). Membranes were scanned and analyzed using the Odyssey Imaging System and software (LI-COR). Alternatively, immunoblotting was performed according to the antibody manufacturer's recommendations using chemiluminescent detection (Perkin Elmer). All western blot images are representative of at least three independent experiments.

Proliferation assays. All assays were performed as previously described by seeding 1,000 cells per well, drug treatments were performed 24 h after seeding and Cell Titer 96 MTS (Promega) was added 72 h later or as described previously[27,30,36]. IL-3 was removed from Ba/F3 cells 48 h before seeding.

Soft agar assays. Anchorage-independent growth was measured by seeding 100,000 cells per well of soft agar in six-well plates as previously described[30]. Medium was changed every 4 d for 2 weeks. Quantification was performed with MetaMorph Offline Version 7.5.0.0 (Molecular Devices).

Fluorescence *in situ* hybridization. FFPE tissue sections were submitted to a dual-color FISH assay using the laboratory developed *NTRK1* break-apart probe (3′ *NTRK1* (SpectrumRed) and 5′ *NTRK1* (SpectrumGreen)). The prehybridization treatment was performed using the reagents from the Vysis Paraffin Kit IV (Abbott Molecular). Hybridization and analysis was performed as previously described[3,30]. Samples were deemed positive for *NTRK1* rearrangement if ≥15% of tumor cells demonstrated an isolated 3′ signal or a separation of 5′ and 3′ signals that was greater than one signal diameter.

siRNA transfection. KM12 cells were transfected with 30 nM *NTRK1* Silencer Select siRNAs (Life Technologies) using siPORT NeoFX transfection reagent (Life Technologies) at 4 μL/mL.

Flow cytometry. Cell-cycle analysis was performed as previously described[3]. Apoptosis was measured using the Vybrant apoptosis YO-PRO/PI kit (Invitrogen). Briefly, KM12 cells were seeded 24 h before treatment at 500,000 cells per well before trypsinization and staining.

Immunohistochemistry. Immunohistochemical studies for thyroid transcription factor-1 (TTF-1) and thyroglobulin were performed using standard procedures to exclude the possibility of a thyroid carcinoma, which can also express TTF-1 (**Supplementary Fig. 20**). Antibody against TTF-1 (Cell Marque, Cat. #CMC-573) was applied at 1:100 dilution and thyroglobulin (Signet, Cat. #228-13) was applied at 1:25 dilution and incubated at 37 °C for 32 min. Detection for TTF-1 was performed using Ventana multiview (UltraView) and detection for thyroglobulin using Ventana Avidin-Biotin (iView).

24. Gnirke, A. *et al. Nat. Biotechnol.* **27**, 182–189 (2009).
25. Li, H. *et al. Bioinformatics* **25**, 2078–2079 (2009).
26. McKenna, A. *et al. Genome Res.* **20**, 1297–1303 (2010).
27. Zhou, W. *et al. Nature* **462**, 1070–1074 (2009).
28. Berger, M.F. *et al. Genome Res.* **20**, 413–427 (2010).
29. Kim, D. *et al. Genome Biol.* **14**, R36 (2013).
30. Doebele, R.C. *et al. Clin. Cancer Res.* **18**, 1472–1482 (2012).
31. Helfrich, B.A. *et al. Clin. Cancer Res.* **12**, 7117–7125 (2006).
32. Davies, K.D. *et al. Clin. Cancer Res.* **18**, 4570–4579 (2012).
33. Marek, L. *et al. Mol. Pharmacol.* **75**, 196–207 (2009).
34. Pitts, T.M. *et al. Front Pharmacol.* **4**, 35 (2013).
35. Oie, H.K. *et al. J. Cell Biochem. Suppl.* **24**, 24–31 (1996).
36. Sasaki, T. *et al. Cancer Res.* **71**, 6051–6060 (2011).

 doi:10.1038/nm.3352